McKean-Vlasov Equations: A Probabilistic and Pathwise Approach

William Salkeld

Doctor of Philosophy
University of Edinburgh
12th May, 2020

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*William Salkeld*)

To my loving family

Acknowledgements

To God, thank you for doing all the hard work.

I am immeasurably grateful to my supervisor Gonçalo dos Reis without whom this whole undertaking would have been impossible. You showed great trust in me from before I applied to work with you and you have had my back completely in the toughest of times. You welcomed me into your family, listened to every one of my personal crises over the last four years and helped to make my present life the joy that it is. I hope that one day I can be as kind and caring a supervisor as you have been to me and the world needs more academics that are as talented as you.

Next, to my parents who have put up with foolish determination to complete this project. Thank you for listening to my rambling explanations over the phone, appeasing my anxiety and keeping me informed with the goings on at home. Also to my sister, I will always be grateful for your companionship. Continue to follow your dreams, you are great at it.

I wish to thank all of my coauthors for their contributions and guidance: Julian Tugaut, Peter Imkeller, Greig Smith, Thomas Cass and Daniel Adams.

To all of my long suffering maths teachers, thank you for having the patience to see me get as far as I have. Special thanks goes to Professor Zhongmin Qian who first introduced me to Malliavin calculus and rough paths.

I am also indebted to all the academics that have helped to encourage my writing with stimulating conversation. In particular, I wish to thank (in no particular order) Peter Friz and Terry Lyons, Annie Millet and Martha Sans Solé, François Delarue, Stefan and Christel Geiss, David Nualart, Mario Maurelli and Zdzislaw Brzezniak, Pierre Cardaliaguet, Andreas Basse-O'Connor, Michael Rochner, Eyal Neuman and Dan Crisan.

For those who couldn't stick around to see this printed on paper, in particular Christabel MacNair and Irene Wells, thank you for the time we had. I hope that I see the world as you did.

To Becky Roberts and those living in the parish of Harnham, thank you for always making sure I have a community to come home to. To the church communities that I have made my home, including St Mary's Longfleet, St Giles Edinburgh and Central Church, thank you for giving me a place to do something other than maths. Special thanks goes to Mae Moore and the members of Living Room who helped me over the line.

To the Brasenose 2011 mathematical matriculands, congratulations guys! We all made it together.

To the writers, producers and cast of The Archers, thank you for 13 minutes of peace six days a week. Shout out to Helen Archer who started the rest of her life as a free woman the day I started my PhD.

To Sara Pascoe, Deborah Frances-White and the many other Guilty Feminists, thank you for helping me better understand myself and my privilege.

Abstract

This thesis divides neatly into four collections of results.

In the first (Part II), we provide existence and uniqueness results along with several properties for a class of McKean-Vlasov Equations having random coefficients and drifts of superlinear growth. We show a Freidlin-Wentzell-type Large Deviations Principles (LDP) in the Hölder topology for the solution process of McKean-Vlasov Equations using techniques that directly address the presence of the law in the coefficients. Our methods avoid using decoupling tricks or particle system approximations.

Secondly (Part III), we close an unexpected gap in the literature concerning the Malliavin and Parametric Differentiability of Stochastic Differential Equations with drifts of super linear growth and with random coefficients. We establish Stochastic Gâteaux Differentiability and Ray Absolute Continuity and take limits in probability rather than mean square or almost surely, bypassing the potentially non-integrable error terms from the unbounded drift.

As an application, we recover representations linking both derivatives as well as a Bismut-Elworthy-Li formula.

Thirdly (Part IV), we prove a representation for the support of McKean-Vlasov Equations. To do so, we construct functional quantizations for the law of Brownian motion as a measure over the (non-reflexive) Banach space of Hölder continuous paths. By solving optimal Karhunen Loève expansions and exploiting the compact embedding of Gaussian measures, we obtain a sequence of deterministic finite supported measures that converge to the law of a Brownian motion with explicit rate. We show the approximation sequence is near optimal with very favourable integrability properties and prove these approximations remain true when the paths are enhanced to rough paths. These results are of independent interest.

The functional quantization results then yield a novel way to build deterministic, finite supported measures that approximate the law of the McKean-Vlasov Equation driven by the Brownian motion which crucially avoid the use of random empirical distributions. These are then used to solve an approximate skeleton process that characterises the support of the McKean-Vlasov Equation.

We give explicit rates of convergence for the deterministic finite supported measures in rough path Hölder metrics and determine the size of the particle system required to accurately estimate the law of McKean-Vlasov equations with respect to the Hölder norm.

Finally (Part V), we study the Small Ball Probabilities of Gaussian rough paths. While many works on rough paths study the Large Deviations Principles for stochastic processes driven by Gaussian rough paths, it is a noticeable gap in the literature that Small Ball Probabilities have not been extended to the rough path framework.

LDPs provide macroscopic information about a measure for establishing Integrability type properties. Small Ball Probabilities provide microscopic information and are used to establish a locally accurate approximation for a random variable. Given the compactness of a Reproducing Kernel Hilbert space ball, its Metric Entropy provides invaluable information on how to approximate the law of a Gaussian rough path. As an application, we are able to find upper and lower bounds for the rate of convergence of an empirical rough Gaussian measure to its true law in pathspace.

Lay summary

"*When you are a Bear of Very Little Brain, and you Think of Things, you find sometimes that a Thing which seemed very Thingish inside you is quite different when it gets out into the open and has other people looking at it.*" - Winnie the Pooh.

In Part II, we study a class of McKean-Vlasov Equations with a one-sided Lipschitz condition for the drift term. McKean-Vlasov Equations model the motion of a particle travelling as part of a cloud of interacting particles. The solution represents the path of the single particle, while the solution law represents the motion of all particles. The one-sided Lipschitz condition means that a particle has multiple stable positions, but if it deviates too far from these then we expect it to be drawn back towards them.

We study the irregularity of these equations. We show that when the driving noise is very small, we expect the solutions to act similarly to the skeleton process of the McKean-Vlasov Equation, a deterministic Ordinary Differential Equation that is not dependent on the law of the McKean-Vlasov.

To prove our results, we compute the probability that the McKean-Vlasov Equation is far from the skeleton process given that the noise remains in a "sausage" centred around a smooth path. Our contributions include accounting for the way in which the law of the McKean-Vlasov Equation acts in these small noise limits, the generality of the assumptions we consider and the ways in which our methods avoid the use of particle systems to approximate the law.

To conclude, we observe that the LDP we have proved implies that a particular scaling phenomena of Brownian motion is also true for McKean-Vlasov equations under the right conditions.

In Part III, we wish to study how much a stochastic process will change when we perturb the solution in different ways. Chapter 9 focuses on the way in which a change in the driving noise affects the solution. The results that we prove are not unexpected, but the most commonly used argument for proving these results fails in an unexpected way. Specifically, we obtain error terms that are non-integrable. Our contribution is the use of a new method that allows us to navigate around these failings of integrability.

Our framework is that of a drift term that satisfies a one-sided Lipschitz condition and random coefficients. The one-sided Lipschitz condition ensures that if the process deviates too far from an attractor, the drift term will push it backwards deterministically. However, the randomness of the coefficients means that the process might deviate far from this attractor at any time. This interaction makes the processes highly unstable. Our results are much more general than those found in standard references such as [Nua06]. We also take care to establish examples that demonstrate the strengths and limitations of our method.

Chapter 10 studies the ways in which a change in the initial condition (both as a fixed constant and as a random variable) affects the solution. The results of both chapters are then combined to prove a Bismut-Elworthy-Li formula for our framework in Chapter 11. This is a standard result in financial mathematics for computing the sensitivities of financial options. Our conclusion is that having random coefficients makes the Bismut-Elworthy-Li formula more involved and previously standard and verifiable assumptions need to be revised.

In Part IV, our goal is to find a set that contains all the paths that we would consider it reasonable for a solution of a McKean-Vlasov Equation to take. This set should be analytic (not probabilistic) and dependent only on $\mathcal{H}$, the drift term b and diffusion term σ. The set $\mathcal{H}$ is the

Reproducing Kernel Hilbert space of the driving noise and contains "all of the information" that we need to understand the driving noise.

There is a classical understanding of this problem. However, this solution fails to be meaningful because in order to understand any one path in the support, one needs to know the law of the McKean-Vlasov Equation. It should not be necessary that this extra piece of information is included, and yet there is no easy way to do without it.

Our solution is to extract extra information from the Reproducing Kernel Hilbert space $\mathcal{H}$ (knowing that the driving noise is Gaussian rather than just a measure with support $\overline{\mathcal{H}}$) to construct a deterministic, finite support measure that approximates the law of the Brownian motion. This is an original approach to the problem and we compare it to the more classical probabilistic method of sampling an empirical distribution. The assembly of this sequence of functional quantizations is explored in Chapter 12.

These finite support measures are used to solve a deterministic system of interacting equations. We use these to obtain a sequence of measures that converge to the law of the McKean-Vlasov Equation. We verify continuity and integrability conditions in Chapter 13 to ensure the new objects that we are assembling are meaningful. Finally, in Chapter 14 we state and prove our expression. In particular, the skeleton process we use to express the support is different from the skeleton process we use to prove the Large Deviations Principle in Chapter 7.

Since their conception by Terry Lyons in [Lyo98] rough paths have become a widely used tool for solving systems of equations being driven by highly oscillating paths. A rough path is an algebraic object that encodes all of the information needed to give a rich theory of integration. This includes more than just the increments of the path, but also the iterated integrals of the path with respect to itself. With this extra information included, it is natural that the focus of attention has been on their macroscopic properties as it is necessary to verify that this information does not render the rough path non-integrable.

Our contribution in Part V is to study the microscopic properties of Gaussian rough paths. Specifically, we compute the asymptotic rate of convergence of the probability that a Gaussian rough path remains in a ball of radius ε.

In Chapter 15, we re-prove several well known (and less well known) results in a new and more general framework. In Chapter 16 we find an expression for the shape of the enhanced Gaussian around 0. Our proof involves developing ideas for the rough path setting and exploiting the compactness properties of Gaussian measures that translate to enhanced Gaussian processes.

Finally, in Chapter 17 we provide some applications to motivate the results of the previous two chapters. In particular, we demonstrate that the compactness properties of the Gaussian measure are translated to the enhanced Gaussian measure, even though the enhancement is not a continuous map. This has applications for the rate of convergence for empirical distributions.

Notation

The following is a list of standard notation that we use throughout this theses:

- $\mathbb{N} = \{1, 2, \cdots\}$, the set of natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$
- $\mathbb{Z}$ and $\mathbb{R}$ the set of integers and real numbers respectively. $\mathbb{R}^+ = [0, \infty)$.
- $\lfloor x \rfloor$ the largest integer less than or equal to $x \in \mathbb{R}$.
- $\mathbb{1}_A$ denotes the usual indicator function over some set A.
- e_j the unit vector of $\mathbb{R}^d$ in the j^{th} component.
- $[0, T]$ a compact, connected, positive lebesgue measure subset of $\mathbb{R}$.
- $C([0, T]; \mathbb{R}^d)$ the space of continuous functions $f : [0, T] \to \mathbb{R}^d$ paired with the supremum norm $\|f\|_\infty = \sup_{t \in [0,T]} |f_t|$.
- $C^\alpha([0, T]; \mathbb{R}^d)$ the space of α-Hölder continuous functions taking values on $\mathbb{R}^d$ paired with the Hölder norm $\|f\|_\alpha = \sup_{s,t \in [0,T]} \frac{|f_t - f_s|}{|t-s|^\alpha}$.
  $C^{\alpha,0}([0, T]; \mathbb{R}^d)$ the closure of $C^\beta([0, T]; \mathbb{R}^d)$ with respect to $\| \cdot \|_\alpha$ where $\beta > \alpha$.
- Let $\tilde{\Omega} = C([0, T]; \mathbb{R}^{d'})$ be the canonical d'-dimensional Wiener space and let W be the Wiener process with law $\tilde{\mathbb{P}}$. Let $(\tilde{\mathcal{F}}_t)_{t \in [0,T]}$ be the standard augmentation of the filtration generated by the Brownian motion. Then we have the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \in [0,T]}, \tilde{\mathbb{P}})$. Additionally, let $([0, 1], \mathcal{B}([0, 1]), \overline{\mathbb{P}})$ be a probability space with the Lebesgue measure $\overline{\mathbb{P}}$. Our probability space is structured as follows:
  1. The sample space will be $\Omega = [0, 1] \times \tilde{\Omega}$,
  2. The σ-algebra over this space will be $\mathcal{F} = \sigma(\mathcal{B}([0, 1]) \times \tilde{\mathcal{F}})$ with filtration $\mathcal{F}_t = \sigma(\mathcal{B}([0, 1]) \times \tilde{\mathcal{F}}_t)$,
  3. The probability measure will be the product measure $\mathbb{P} = \overline{\mathbb{P}} \times \tilde{\mathbb{P}}$.
- For $p \geq 1$, let $L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ be the space of random variables over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with state space $\mathbb{R}^d$ and finite p moments.
- Let $\mathcal{S}^p([0, T]; \mathbb{R}^d)$ be the space of random processes such that $\mathbb{E}\Big[\|X\|_\infty^p\Big] < \infty$ where $X : [0, T] \times \Omega \to \mathbb{R}^d$.
- $\mathcal{H}$ be the Reproducing Kernel Hilbert space, sometimes referred to as the Cameron Martin space of an associated Gaussian process (often but not exclusively Brownian motion).
- $f_n \lesssim g_n \iff \limsup_{n\to\infty} \frac{f_n}{g_n} \leq C$ and $f_n \gtrsim g_n \iff \liminf_{n\to\infty} \frac{f_n}{g_n} \geq C$ for sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$.
  When $f_n \lesssim g_n$ and $f_n \gtrsim g_n$, we say $f_n \approx g_n$. In contrast $f_n \sim g_n \iff \lim_{n\to\infty} \frac{f_n}{g_n} = 1$.
- $\nabla_x f$ the gradient of a differentiable function f.

Contents

Part I

Introduction

Summary

This thesis is the compilation of two published papers ([dRST19] and [IdRS19]) and two preprints ([CRS19] and [Sal20]). Thus, it naturally divides into four parts:

Part II has three main results. Firstly, in Chapter 6 we prove sharp existence and uniqueness result for McKean-Vlasov Equations with a one-sided Lipschitz drift term and random drift and diffusion terms. In Chapter 7 we prove Large Deviations Principle for McKean-Vlasov Equations and apply this in Chapter 8 to obtain a Functional Iterated Logarithm law for a class of McKean-Vlasov Equations.

While writing [dRST19], we believed that it would be an easy extension of existing results to prove Malliavin differentiability of the McKean-Vlasov Equations considered. This turned out to be more challenging that we expected due to the gap in the literature surrounding differentiability of SDEs with a one-sided Lipschitz drift term. This was the inspiration for writing [IdRS19]. Part II proves Malliavin Differentiability in 9 and Parametric Differentiability in 10 for a wide class of stochastic differential equation. Applications are discussed in 11.

Additionally, while writing [dRST19] we believed it would be a standard application of the LDP results to obtain a support theorem for McKean-Vlasov Equations. This is the case for classical SDEs, but the presence of the law in the skeleton process complicates matters. Thus the goal of Part IV is to prove a Support theorem for McKean-Vlasov Equations. To do this, we solve a quantization for the driving Gaussian noise in Chapter 12. Properties of particle systems driven by quantizations and their integrability is discussed and proved in Chapter 13 and finally the representation of the support is proved in Chapter 14.

Part V takes a slightly different approach. The project started while researching the scope of [CRS19] and establishing the difference between quantization of Gaussian measures over Banach spaces and enhanced Gaussian measures over the collection of rough paths. The goal is to study the rate of convergence of empirical distributions of enhanced Gaussian white noises to the true enhanced Gaussian measure. To do this, we demonstrate several Gaussian correlation inequalities in Chapter 15. These are applied to compute the small ball probabilities of Gaussian rough paths in Chapter 16. Finally, these are applied in Chapter 17 to compute the strong rate of convergence of the empirical distribution to the true law of an enhanced Gaussian measure via its metric entropy.

Chapter 1

McKean-Vlasov Equations

This thesis primarily focuses on the study of McKean-Vlasov Equations (MVEs). McKean-Vlasov Equations are Stochastic Differential Equations where the coefficients of the equation are dependent on the law of the solution. Typically, they are denoted

$$dX_t = b(t, X_t, \mathcal{L}_t^X)dt + \sigma(t, X_t, \mathcal{L}_t^X)dW_t, \quad X_0 = \theta \tag{1.0.1}$$

where $\mathcal{L}_t^X$ is the pushforward measure of the solution $\mathcal{L}_t^X = \mathbb{P} \circ (X_t)^{-1}$ and $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space containing a Gaussian white noise that will typically be Brownian motion. These equations were first studied by [McK66] in the context of statistical physics and developing the ideas of [Kac56] to give meaning to a system of weakly interacting particles of the form

$$dX_t^{i,N} = b\Big(t, X_t^{i,N}, \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}\Big)dt + \sigma\Big(t, X_t^{i,N}, \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}\Big)dW_t^{i,N}, \quad X_0^{i,N} = \theta^i \tag{1.0.2}$$

where $i \in \{1, ..., N\}$, each W^i is an independent Gaussian noise and θ^i are independent, identically distributed random variables independent of the Brownian motions. Thus each particle in this "cloud of particles" is dependent on the position of any other particle but as N becomes large the equations are increasingly determined by the distribution of the cloud of particles rather than the position of any one particle. As N becomes large, Equations of the form (1.0.2) become computationally challenging to solve due to the cost of simulating large numbers of independent white noises. However, the motion of a single equation within the system of interacting equations can be approximated by (1.0.1) as $N \to \infty$ using the well known results of [Szn91], see also [Mél96] and [Car16], which states

$$\lim_{N\to\infty} \sup_{1\leq i\leq N} \mathbb{E}\Big[\|X^{i,N} - X^i\|_\infty^2\Big] = 0. \tag{1.0.3}$$

Generally, such results are referred to as *Propagation of chaos* and imply that the dynamics of Equation (1.0.2) can be approximated by (1.0.1) when N is large. However, the study of Equation (1.0.1) is also of independent interest

Since their conception, McKean-Vlasov Equations have proved a powerful tool for modelling in a wide variety of fields. Applications are numerous and vary from opinion dynamics [HK02], the dynamics of granular materials [BCCP98, BGG13, CGM08], molecular and fluid dynamics [Pop01], interacting agents in economics or social networks [CDL13], mathematical biology [KS71, BCM07], Galactic dynamics [BT11], droplet growth[CS19], Plasma Physics [Bit13], interacting neurons [DIRT15] and deep learning neural networks [HKR19]. See [CD17a, CD17b] and references therein for a detailed exploration of the applications of McKean-Vlasov Equations.

1.1 The space of measures

For E a complete, separable metric space with Borel σ-algebra $\mathcal{B}$, let $\mathcal{P}_r(E)$ be the set of all Borel measures over $(E, \mathcal{B})$ which have finite r^{th} moments.

Definition 1.1.1. *Let $r \geq 1$. Let $\mu, \nu \in \mathcal{P}_r(E)$. We define the Wasserstein r-distance $\mathbb{W}^{(r)}_{E,d} : \mathcal{P}_r(E) \times \mathcal{P}_r(E) \to \mathbb{R}^+$ to be*

$$\mathbb{W}^{(r)}_{E,d}(\mu, \nu) = \left(\inf_{\gamma \in \mathcal{P}(E\times E)} \int_{E \times E} d(x,y)^r \gamma(dx, dy) \right)^{\frac{1}{r}} \tag{1.1.1}$$

where γ is a joint distribution over $E \times E$ which has marginals μ and ν. When the space the measure is defined on is clear, we write $\mathbb{W}^{(r)}_d$ where d is the metric over E.

The problem of finding a measure $\gamma \in \mathcal{P}_2(E \times E)$ that minimises (1.1.1) is sometimes referred to as the Kantorovich problem and γ is called the transport plan of μ and ν. The choice of $r = 2$ is common throughout literature. However, we will also be interested in the case $r = 1$. Sometimes referred to as the Kantorovich–Rubinstein metric, the Wasserstein distance induces the topology of weak convergence of measure as well as convergence in moments of order up to and including r. The Wasserstein distance is a metric, but the metric does not induce a norm as it is homogeneous but not translation invariant. The space $\mathcal{P}_2(E)$ is complete and separable with respect to the Wasserstein metric (see [Bol08]).

It is important to define the Wasserstein distance for a generic complete separable metric space because later on we will be interchanging between measures on $\mathbb{R}^d$ and $C([0,T];\mathbb{R}^d)$. In order to distinguish between measures on $\mathbb{R}^d$ and $C([0,T];\mathbb{R}^d)$, we denote $\mu \in \mathcal{P}_2(C([0,T];\mathbb{R}^d))$ and $\mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ and we define for $A \subset \mathbb{R}^d$

$$\mu_t(A) = \int_{C([0,T];\mathbb{R}^d)} \mathbb{1}_{\left\{x(\cdot) \in C([0,T];\mathbb{R}^d);\, x(t) \in A\right\}}(x) \mu(dx).$$

For an in-depth treatment of Optimal transport, we refer the reader to [Vil09] or [CD17a, Chapter 5].

1.1.1 Examples of measure dependencies

Given complete, separable metric spaces E and F and a function $f : \mathcal{P}_2(E) \to F$, it is natural to ask what forms f might take.

Example 1.1.2 (Scalar interactions)**.** *Let $\phi : E \to \mathbb{R}^d$ be a Lipschitz function, let $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and let $\mu \in \mathcal{P}_1(E)$. Let*

$$f(x, \mu) := F\Big(x, \int_E \phi d\mu\Big).$$

We call f a mean-field interaction of scalar type.

Example 1.1.3 (Interactions of order 1 and beyond)**.** *Let $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function with constant L and let $\mu \in \mathcal{P}_1(\mathbb{R}^d)$. Let*

$$f(x, \mu) := \int_{\mathbb{R}^d} F(x,y) \mu(dy).$$

We call f a mean-field interaction of order 1. Observe that for $\mu, \nu \in \mathcal{P}_1(E)$

$$|f(x,\mu) - f(x,\nu)| = \Big| \int_{\mathbb{R}^d} F(x,y)\mu(dy) - \int_{\mathbb{R}^d} F(x,z)\mu(dz) \Big| \leq L \cdot \int_{\mathbb{R}^d \times \mathbb{R}^d} |y - z| \gamma(dy, dz)$$

where γ has marginals μ and ν. Minimising over the choice of γ, we get

$$|f(x,\mu) - f(x,\nu)| \leq L \cdot \mathbb{W}^{(1)}_{\mathbb{R}^d}(\mu, \nu).$$

Similarly, let $F : (\mathbb{R}^d)^{\times(n+1)} \to \mathbb{R}$ be a Lipschitz function and let $\mu \in \mathcal{P}_1(\mathbb{R}^d)$. Then we say that the functional

$$f(x,\mu) := \int_{\mathbb{R}^d} ... \int_{\mathbb{R}^d} F(x, y_1, ..., y_n)\mu(dy_1)...\mu(dy_n)$$

is a mean-field interaction of order n.

1.2 Existence and Uniqueness

We start with a slight generalization of the existence and uniqueness result under Lipschitz conditions in [Car16, Theorem 1.7]. Let W be a d'-dimensional Brownian motion and let $b : [0,T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d'}$ be progressively measurable maps.

For $0 \le t \le T < \infty$, we introduce the dynamics of a process Y

$$dY_t = b\big(t, Y_t, \mathcal{L}^Y_t\big)dt + \sigma\big(t, Y_t, \mathcal{L}^Y_t\big)dW_t, \tag{1.2.1}$$

where $Y_0 \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$ and $\mathcal{L}^Y_t$ denotes the Law of Y_t.

Theorem 1.2.1. *Suppose that b and σ are integrable in the sense that*

$$\mathbb{E}\Big[\Big(\int_0^T |b(t,0,\delta_0)|dt\Big)^2\Big] < \infty \quad \textit{and} \quad \mathbb{E}\Big[\int_0^T |\sigma(t,0,\delta_0)|^2 dt\Big] < \infty,$$

and Lipschitz in the sense that $\exists L > 0$ such that for almost all $(t,\omega) \in [0,T] \times \Omega$, $\forall x, x' \in \mathbb{R}^d$ and $\forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have that

$$\begin{aligned}|b(t,\omega,x,\mu) - b(t,\omega,x',\mu')| + |\sigma(t,\omega,x,\mu) - \sigma(t,\omega,x',\mu')|& \\ \le L\Big(|x-x'| + \mathbb{W}^{(2)}(\mu,\mu')\Big).&\end{aligned}$$

Suppose further that $Y_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ is a square integrable random variable which is independent of the Brownian motion. Then there exists a unique solution for $Y \in \mathcal{S}^2([0,T]; \mathbb{R}^d)$ to the McKean-Vlasov Equation (1.2.1) *and $\mathcal{L}^Y_0 \in \mathcal{P}_2(\mathbb{R}^d)$ where $\mathcal{L}^Y_t$ is the probability distribution of the random variable Y_t.*

Proof. For $b(\cdot, 0, \delta_0)$ satisfying $\mathbb{E}[\int_0^T |b(t,\omega,0,\delta_0)|^2 dt] < \infty$ the result is known e.g. [Car16, Theorem 1.7]. A close inspection of that proof shows that this condition is not sharp. In particular, the result holds with the slightly weaker integrability condition found in the statement of the theorem we present here. The verification is straightforward and we do not carry it out. $\square$

Historically, the existence problem for (1.2.1) has been investigated by two different methods. The first one consists in the application of a fixed point theorem on the space of measures, see [McK67], or on a collection of permissible drifts, see [BRTV98, CGM08, HIP08]. The other method consists of proving the law of the associated particle system converges to the solution of a non-linear Martingale problem, see for example [Mél96, BJT11, FJ$^+$17] .

More recently, Zvonkin's transformation tools were extended to the mean-field setting in [dR20] to allow for drift terms that are only Hölder continous in the spacial and measure dependencies provided the diffusion term is non-degenerate. Similarly, [MV16] was able to prove Existence and Uniqueness assuming that the drift has linear growth in the spacial dependency and is bounded in the measure dependency, and a non-degeneracy of the diffusion. These ideas were later extended in [HŠS18] to prove a weak existence for McKean-Vlasov Equations with unbounded coefficients that satisfy a Lyapunov condition. Further results on existence and uniqueness can be found in [Car16, CD17a, CD17b]. We highlight [Sch87] for a discussion on counterexamples on uniqueness of solutions.

1.3 Our contributions

In Chapter 6, we prove an existence and uniqueness result for McKean-Vlasov Equations with a drift terms satisfying a one-sided Lipschitz in the spacial variable (see Theorem 6.1.2). Further, the coefficients are random and dependent on time, and both drift and diffusion terms are Lipschitz in the measure variables. We prove our results by dealing with the presence of the laws in the coefficients directly and avoiding arguments on empirical measures or approximation/-convergence of measures.

Our method for proving existence and uniqueness uses stopping time arguments, but critically it is not the McKean-Vlasov Equation that is stopped but an SDE with an approximation of the law substituted into the coefficients. This is an important detail because stopped McKean-Vlasov Equations are not themselves McKean-Vlasov: the law of the solution is not the same as the law in the coefficients. Further, we provide sharp integrability conditions for the (t,ω) dependency in the drift and diffusion terms in Assumption 6.1.1.

Finally, in Section 6.2 we demonstrate the regularity in time of these processes. In particular, we observe that the addition of measure dependency does not complicate the regularity as it is generally smoother than the contributions from the noise.

Chapter 2

Large Deviations Principle and Applications

The goal of a Large Deviations Principle (LDP) is to find asymptotic upper and lower bounds for the probability of a rare event. The problem was first studied by Cramér and Lundberg to quantify the vulnerability of insurers to insolvency and was later developed for stochastic differential equations in [Var66]. For this chapter, we briefly outline the basic ideas for Large Deviations Principles, but for a more detailed exploration of the field see [DE97] and [DZ10].

The large deviations principle characterises the limiting behaviour of a family of probability measures $(\mathcal{L}_\varepsilon)_{\varepsilon>0})$ over $(E, \mathcal{B})$, where E is a topological space and $\mathcal{B}$ is the Borel σ-algebra, using a rate function $\mathcal{I}$.

Definition 2.0.1. *A rate function $\mathcal{I}$ is a lower semicontinous mapping $\mathcal{I} : E \to [0, \infty]$ such that $\forall a > 0$, the level sets $\{\phi : \mathcal{I}(\phi) < a\}$ are closed. A rate function is called a good rate function if $\forall a > 0$ the level sets are compact.*

Given a Borel set A, we characterise the limiting behaviour of the sequence of measures in terms of the rate function restricted to the interior and exterior of A.

Definition 2.0.2. *The collection of measures $(\mathcal{L}^\varepsilon)_{\varepsilon>0}$ satisfies a* Large Deviations Principle *with rate function $\mathcal{I}$ if and only if for any Borel set A, we have*

$$-\inf_{\phi\in A^\circ} \mathcal{I}(\phi) \le \liminf_{\varepsilon\to 0} \varepsilon \log\Big(\mathcal{L}^\varepsilon[A]\Big) \le \limsup_{\varepsilon\to 0} \varepsilon \log\Big(\mathcal{L}^\varepsilon[A]\Big) \le -\inf_{\phi\in \overline{A}} \mathcal{I}(\phi).$$

With regard to stochastic processes such as McKean-Vlasov Equations, we endeavour to establish a deterministic path around which the diffusion is concentrated with high probability. The consequence of this is that we can think of the McKean-Vlasov Equation as a small probabilistic perturbation of an Ordinary Differential Equation.

Consider the following motivational example:

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}dW_t, \quad X_0 = x_0, \tag{2.0.1}$$

where $x_0 \in \mathbb{R}^d$, $b : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz and W is a d'-dimensional Brownian motion. We pair this stochastic differential equation with the operator from the Cameron Martin space $\mathcal{H}$ to the deterministic ODE

$$d\Phi[h]_t = b(\Phi[h]_t)dt, \quad \Phi[h]_0 = x_0. \tag{2.0.2}$$

Let $\mathcal{L}^\varepsilon = \mathbb{P}\circ(X^\varepsilon)^{-1}$ be the law of Equation (2.0.1) over the Banach space of continuous functions equipped with the supremum norm $\|\cdot\|_\infty$.

It is well known that $\mathcal{L}^\varepsilon$ satisfies a Freidlin-Wentzell LDP with respect to the supremum norm with rate function

$$\mathcal{I}\Big(\Phi[h]\Big) = \frac{\|h\|_{\mathcal{H}}^2}{2}.$$

In this example, the rate function is defined explicitly to be

$$\mathcal{I}(\phi) := \begin{cases} \frac{1}{2}\int_0^T \|\dot{\phi}_t - b(\phi_t)\|^2 dt & \iff \phi \text{ is absolutely continuous,} \\ \infty & \text{otherwise.} \end{cases}$$

2.1 Recent LDP results for McKean-Vlasov Equations

The LDP results we present are with respect to the vanishing noise as in Freidlin-Wentzell theory, see [FW12]. For instance, [HIP08] investigates the large deviations principle for the McKean-Vlasov diffusion

$$X_t^\varepsilon = X_0 - \int_0^t \nabla_x V\big(X_s^\varepsilon\big) ds - \int_0^t \nabla_x F * u_s^\varepsilon\left(X_s^\varepsilon\right) ds + \sqrt{\varepsilon} W_t, \tag{2.1.1}$$

$$\varphi(t) = X_0 - \int_0^t \nabla_x V(\varphi(s)) ds \tag{2.1.2}$$

in general dimensions, assuming superlinear growth of the drift but imposing coercivity in their monotonicity condition and a constant diffusion term. In particular, they show that the family of laws $(\nu_\varepsilon)_\varepsilon$ satisfies a large deviations principle on $C([0,T];\mathbb{R}^d)$ equipped with the uniform norm with the good rate function

$$I(\phi) := \frac{1}{2}\int_0^1 \left\|\dot{\phi}(t) + \nabla_x V\left(\phi(t)\right) + \nabla_x F\big(\phi(t) - \varphi(t)\big)\right\|_\infty^2 dt\,,$$

when ϕ is absolutely continuous such that $\phi(0) = X_0$ and $I(\phi) := +\infty$ otherwise (φ as in (2.1.2)).

A common way in which Large Deviation Principles are used to study McKean-Vlasov Equations is via Propagation of Chaos and the so called "Sanov's theorem". First introduced in [San61], these type of results explore the rate of convergence of the empirical measure to the true law of the McKean-Vlasov Equation. While it is true that these are similar to a Freidlin Wentzell type result in that they study the convergence in probability of random distributions over the space of continuous paths, they are distinct and of independent interest from the Freidlin-Wentzell results we study in this work.

The sutdy of Sanov-type results is a huge field of research and we highlight only a small number of relevant works [BDF12, DPdH96, DG87, DFMS18] and the references within.

2.2 Our Contributions

We show a similar Large Deviations Principle results in the Hölder norm, for the family of measures associated to

$$dX_t^\varepsilon = b_\varepsilon\Big(t, X_t^\varepsilon, \mathcal{L}_t^{X^\varepsilon}\Big)dt + \sqrt{\varepsilon}\sigma_\varepsilon\Big(t, X_t^\varepsilon, \mathcal{L}_t^{X^\varepsilon}\Big)dW_t. \tag{2.2.1}$$

However, unlike [HIP08], we assume a Lipschitz σ coefficient (not a constant one) and we do not impose any coercivity condition (strict negativity of the monotonicity constant).

We draw inspiration from [BAL94]. Studying standard SDEs, the authors find a way to transfer LDP results from a coarse topology to a finer one; in their case, from supremum norms to Hölder norms. Their method, explained later, relies on establishing the following inequality: $\forall R > 0, \forall \rho > 0$, $\exists \delta > 0$ and for ε small enough (see Theorem 7.2.3 below for the precise statement)

$$\mathbb{P}\Big[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta\Big] \lesssim \exp\Big(-\frac{R}{\varepsilon}\Big), \tag{2.2.2}$$

for classical SDE's where $\Phi^x(h)$ is the so-called skeleton map (an ODE) associated with X_ε^X. This can be thought of as establishing that the probability of X having a high variation in the $\|\cdot\|_\alpha$-norm given that the input signal (from the Brownian motion) is small in $\|\cdot\|_\infty$-norm is exponentially small. For this, they assume boundedness and Lipschitz properties of the drift and diffusion coefficients of the SDE X^ε dependent only on the spatial variables. We provide results in the same vein but for the general class of McKean-Vlasov Equations with drifts of polynomial growth (see Assumption 7.2.1). Their conditions are stronger than our conditions so our results extend existing results in classical SDE literature.

In [dRST19], we additionally prove Large Deviations Principle results for McKean-Vlasov Equations of the form (2.2.1) with respect to the supremum. These results are not included in this thesis as they were collaborative results with Dr. Tugaut but are key to our methods for proving Proposition 7.2.3.

Our results on LDPs are of general interest and can be applied to the Monte-Carlo simulation of McKean-Vlasov Equations in the spirit of [GR08] as a way to find the optimal Importance Sampling measure, see [dRST18]. Applications of our results also include Chapter 8 and simulation of McKean-Vlasov Equations (see [dRES18] and [RW18]).

Chapter 3

Malliavin Calculus

Malliavin calculus, or Stochastic calculus of variations, is a collection of tools first developed by Paul Malliavin in the 1970's to study the densities of stochastic processes. It relies on the powerful structures of Gaussian measures and the ways in which these underpin the theory of stochastic processes today. In this chapter, we provide a short introduction to Gaussian measures and Malliavin calculus. For an overview of Gaussian analysis and probability theory on Banach spaces, see [LT13, Bog98]. For detailed references on Malliavin calculus, see [Nua06, DNØkP09, DP06].

3.1 Motivation and background

Let E be a Banach space and let $\mathcal{B}$ be the cylinder σ-algebra, the smallest σ-algebra such that every element of the dual space E^* is measurable. When E is separable, $\mathcal{B}$ is known to be equivalent to the Borel σ-algebra. We say that $\mathcal{L}$ is a Gaussian measure on $(E,\mathcal{B})$ if $\forall f \in E^*$ the pushforward of the measure $\mathcal{L}$ by the linear map f is normally distributed. Thus for each Gaussian measure on $(E,\mathcal{B})$, we have a symmetric Bilinear form $\mathcal{R} : E^* \times E^* \to \mathbb{R}$ defined by $\mathcal{R}(f_1, f_2) := \int_E f_1(x) \cdot f_2(x) d\mathcal{L}(x)$ which we call the covariance operator. Associated to the covariance operator is the covariance kernel $\mathcal{S} : E^* \to E$ defined by the pettis integral

$$\mathcal{S}[f] = \int_E f(x) \cdot x\mathcal{L}(dx).$$

Thus $f(\mathcal{S}[g]) = \mathcal{R}(f,g)$. The Kernel is a positive, self adjoint operator so has a spectral representation

$$\mathfrak{i} : E^* \to L^2(E,\mathcal{L};\mathbb{R}), \quad \mathfrak{i}^* : L^2(E,\mathcal{L};\mathbb{R}) \to E, \quad \mathfrak{i}^*\mathfrak{i} = \mathcal{S}$$

where $\mathfrak{i}$ canonically embeds the linear functionals into the space of square-integrable functionals over E with respect to the measure $\mathcal{L}$. The Reproducing Kernel Hilbert space (RKHS) of the Gaussian measure $\mathcal{L}$ is the completion of the range $S[E^*]$ with respect to the inner product

$$\Big\langle S[f_1], S[f_2] \Big\rangle_{\mathcal{H}} := \int_E f_1(x) f_2(x) \mathcal{L}(dx).$$

The support of the Gaussian meaure $\mathcal{L}$ over $(E,\mathcal{B})$ is the closure of $\mathcal{H}$ is the topology of E. When $\mathcal{H}$ is not dense in E, we call $\mathcal{L}$ degenerate.

Let E be a separable Banach space, let $\mathcal{H}$ be a separable Hilbert space and let $\mathfrak{i}^* : \mathcal{H} \to E$ be an injective continuous linear map with dense image that radonifies the canonical cylinder set measure. Then we say $(E,\mathcal{H},\mathfrak{i}^*)$ is an abstract Wiener space.

Let E be a separable Banach space. A mapping $f : E \to \mathbb{R}$ is called a Polynomial functional of E if $\exists n \in \mathbb{N}, \exists g_1, ..., g_n \in E^*$ and $\exists \tilde{f} : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x) = \tilde{f}(g_1(x), ..., g_n(x)).$$

The set of all polynomial functionals of E is denoted $\mathbf{P}(E)$. Similarly, for separable Banach spaces E_1 and E_2 a mapping $F : E_1 \to E_2$ is called a E_2 polynomial of E_1 if $\exists m \in \mathbb{N}, \exists f_1, ..., f_m \in \mathbf{P}(E_1)$

and $\exists e_1, ..., e_m \in E_2$ such that

$$F(x) = \sum_{j=1}^{m} f_j(x) b_j.$$

The set of all E_2 valued polynomial functionals of E_1 is denoted $\mathbf{P}(E_1, E_2)$.

It is well known that the space $\mathbf{P}(E_1, E_2)$ is dense in the space $L^p(E_1, \mathcal{B}, \mathcal{L}; E_2)$, see for example [Sug85].

For $F \in \mathbf{P}(E_1, E_2)$, we define the derivative of F, denoted DF, to be the $E_2 \otimes \mathcal{H}$ valued polynomial functional of E_1 defined by

$$DF(x) = \sum_{k=1}^{n} \sum_{j=1}^{m} \partial_{x_k} f_j(g_1(x), ..., g_n(x)) e_j \otimes (g_k \circ \mathfrak{i}^*).$$

The operator D is closable, and we define the Sobolev space $\mathbb{D}^{1,p}(E_2)$ to be closure of $\mathbf{P}(E_1, E_2)$ with respect to the norm

$$\|F\|_{1,p,E_2} \left[\int_{E_1} \|F(x)\|_{E_2}^p \mathcal{L}(dx) + \int_{E_1} \Big\| \|DF(x)\|_{\mathcal{H}} \Big\|_{E_2}^p \mathcal{L}(dx) \right]^{1/p}.$$

3.2 Overview of the methodology

It is important to note that the solution of an SDE is not continuous with respect to $\omega \in \Omega$. As the SDE exists in a probability space with the filtration generated by an d'-dimensional Brownian motion, ω can be interpreted to mean the path of an individual Brownian motion plus any extra information about what happens when $t = 0$. However, it will be shown that the random variables are continuous, and indeed differentiable, when perturbed with respect to a path out of the Cameron Martin space. Hence for this section we take $h \in \mathcal{H}$, the Cameron Martin space of a d'-dimensional Brownian motion.

We start by introducing the concepts of *Ray absolute continuity* and *Stochastic Gâteaux Differentiability* and the results yielding Malliavin differentiability under those properties.

Let E be a separable Banach space. Let $L(\mathcal{H}, E)$ be the space of all bounded linear operators $V : \mathcal{H} \to E$.

Definition 3.2.1 (Ray Absolutely Continuous map)**.** *A measurable map $f : \Omega \to E$ is said to be* Ray Absolutely Continuous *if $\forall h \in \mathcal{H}$, $\exists$ a measurable mapping $\tilde{f}_h : \Omega \to E$ such that*

$$\tilde{f}_h(\omega) = f(\omega) \quad \mathbb{P}\text{-}a.e.$$

and that $\forall \omega \in \Omega$,

$$t \mapsto \tilde{f}_h(\omega + th) \quad \textit{is absolutely continuous on any compact subset of } \mathbb{R}.$$

Definition 3.2.2 (Stochastically Gâteaux differentiable)**.** *A measurable mapping $f : \Omega \to E$ is said to be* Stochastically Gâteaux differentiable *if there exists a measurable mapping $F : \Omega \to L(\mathcal{H}, E)$ such that $\forall h \in \mathcal{H}$,*

$$\frac{f(\omega + \varepsilon h) - f(\omega)}{\varepsilon} \xrightarrow{\mathbb{P}} F(\omega)[h] \quad \textit{as } \varepsilon \to 0.$$

Malliavin differentiability follows from [Sug85, Theorem 3.1] which was later improved upon by [MPR17, Theorem 4.1]. We recall both results next.

Theorem 3.2.3 ([Sug85])**.** *Let $p > 1$. The space $\mathbb{D}^{1,p}(E)$ is equivalent to the space of all random variables $f : \Omega \to E$ such that $f \in L^p(\Omega; E)$ is Ray Absolutely Continuous, Stochastically Gâteaux differentiable and the Stochastic Gâteaux derivative $F : \Omega \to L(\mathcal{H}, E)$ is $F \in L^p(\Omega; L(\mathcal{H}, E))$.*

Remark 3.2.4. *We know from standard references such as* [ÜZ00] *that the map $t \mapsto \tilde{f}_h(\omega + th)$ is continuous as a map from $[0, 1] \to L^0(\Omega)$. The point of proving the stronger absolute continuity is* *to find a representation of the form*

$$\tilde{f}_h(\omega + \varepsilon h) - \tilde{f}_h(\omega) = \int_0^\varepsilon F(\omega + rh)[h]dr,$$

where the object $F(\omega)$ is a candidate for the Malliavin Derivative. Proving Stochastic Gâteaux Differentiability is then verifying that this object is a bounded linear operator and allows one to extend from Gâteaux to Fréchet. Thus a random variable which is Ray Absolutely Continuous but not Stochastic Gâteaux Differentiable has a Malliavin Directional Derivative in all directions, but there is a sequence of elements $h_n \in \mathcal{H}$ such that $F(\omega)[h_n] \to \infty$.

By contrast, if one has Stochastic Gâteaux Differentiability but not Ray Absolute Continuity, then one can prove existence of the Malliavin Derivative but which is not in $L^1(\Omega)$ e.g. $\mathbb{E}[\|F(\omega)\|_{L(\mathcal{H},E)}] = \infty$.

3.3 Recent works

The following definition was first introduced in [MPR17].

Definition 3.3.1 (Strong Stochastically Gâteaux differentiable)**.** *Let $p > 1$. A random variable $f \in L^p(\Omega; E)$ is said to be* Strong Stochastically Gâteaux differentiable *if there exists a measurable mapping $F : \Omega \to L(\mathcal{H}, E)$ such that $\forall h \in \mathcal{H}$*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left\|\frac{f(\omega + \varepsilon h) - f(\omega)}{\varepsilon} - F(\omega)[h]\right\|\right] \to 0. \tag{3.3.1}$$

Theorem 3.3.2 ([MPR17])**.** *Let $p > 1$. The space $\mathbb{D}^{1,p}(E)$ is equivalent to the space of all random variables $f \in L^p(\Omega; E)$ that are Strong Stochastically Gâteaux differentiable and have measurable mapping $F \in L^p(\Omega; L(\mathcal{H}, E))$.*

The merit of [Sug85] is that it allows one to prove Malliavin differentiability by first establishing existence of a Gâteaux derivative and then extending to the full Frechét derivative. The convergence of the Gâteaux derivative in probability is a very weak condition that is much easier to prove than full Malliavin differentiability.

[MPR17] extends this result to the stronger Strong Stochastic Gâteaux Differentiability condition and removed the Ray Absolute Continuity condition.

Both of these methods have their merits. While studying different examples of processes with monotone growth, we became interested in the particular example where the drift term has polynomial growth of order q but only finite moments up to $p < q - 2$. In this case, one cannot in general find a dominating function for the error terms coming from the drift of the SDE while trying to prove Stochastic Gâteaux Differentiability. It therefore became necessary to prove only a convergence in probability statement.

The question of Malliavin differentiability of stochastic processes with coefficients that are themselves dependent on the Gaussian noise is somewhat technical and has often been avoided in previous works. In principle, one would expect an additional assumption of Malliavin differentiability of the coefficients plus a joint continuity condition with spacial variables. One work that does address the Malliavin Differentiability of Lévy driven Backwards Stochastic Differential Equations, [GS16], includes adequate regularity conditions for the coefficients to ensure the solution process is Malliavin differentiable.

3.4 Our Contributions

We work with the class of Stochastic Differential Equations (SDEs) with drifts satisfying a superlinear growth (locally Lipschitz) and a monotonicity condition (also called one-sided Lipschitz condition); further, the coefficients are assumed to be random. This class of SDEs appears ubiquitously in mathematics and engineering, for example, the stochastic Ginzburg-Landau equation in the theory of superconductivity; Stochastic Verhulst equation; Feller diffusion with logistic growth; Protein Kinetics and others, see [HJK11] and references.

There is a wealth of results on differentiability and properties of SDEs in general. However, it is surprising that the landscape is (to the best of our knowledge) sparse with respect to the superlinear growth setting apart from [TZ15] which we discuss below. Additionally, in [RS17] the authors discuss stochastic flows in rough path sense for a class related to ours but only up to linear growth; and using analytical tools, [Cer01, Chapter 1] and [Zha16] require ellipticity and deterministic maps to obtain some results in the same vein as ours. Our arguments are fully probabilistic.

To establish Malliavin differentiability for an SDE with solution X and with monotone drifts, the most natural path to follow is to try to apply [Nua06, Lemma 1.2.3] by employing a truncation procedure. This yields a sequence X^n of SDEs with Lipschitz coefficients converging to X. Under said Lipschitz conditions the family X^n is Malliavin differentiable under suitable differentiability assumptions, with derivative DX^n, and one is able to appeal to [Nua06, Lemma 1.2.3] to conclude the Malliavin differentiability of X if one is able to show that $\sup_n \mathbb{E}\big[\|DX^n\|_H\big] < \infty$. The truncation procedure, even smoothed out, destroys the monotonicity and, in the multi-dimensional case, it is notoriously difficult to establish the aforementioned uniform bound.

To the best of our knowledge this question was studied only in [TZ15]. The authors employ a truncation procedure in order to use [Nua06, Lemma 1.2.3]. Unfortunately their [TZ15, Lemma 4.1] is incorrect. The constant M_l presented in their equation (4.1) depends on the truncation level n in a non-uniformly bounded way; the reader is invited to inspect the 2nd line of page 879. This lemma, which we were not able to fix, is used subsequently to establish the main result in [TZ15].

We prove Malliavin Differentiability through a less well-known method developed by Sugita [Sug85] which uses the concepts of *Ray Absolute Continuity* and *Stochastic Gâteaux Differentiability* see also the posterior developments by [MPR17, IMPR16]. This approach is detailed in Section 3.2 above. The merit of this method is that the limit for the Stochastic Gâteaux derivative is a convergence in probability statement rather than a convergence in mean square statement. Put simply, this allows us to avoid cases such as the “Witches Hat” function where errors are non-integrable but converge to zero almost surely.

We study the case where the coefficients of the SDE are random. We follow the ideas of [GS16] and present two different sets of conditions which allow for Malliavin Differentiability. One set of conditions is sharp but somewhat difficult to use in practice. The other is much easier to verify but not sharp. We also provide examples discussing the scope and limitations of our approach.

In this setting the drift term is not bounded and, conditional on the coefficients' integrability, the solution may not be sufficiently integrable - see Remark 9.1.3 and the examples in Section 9.3.3. This means that the error terms appearing in the proofs of differentiability will not be assumed to be sufficiently integrable. We negotiate this obstacle by proving everything in convergence in probability and ensuring that adequate conditions are met so that results can be lifted to the relevant setting of mean square and almost sure convergence.

Chapter 4

Support Theorems

The support of a measure is the smallest closed set of full measure. Thus the *Support theorem* for the law of an SDE characterises the set of admissible paths that the SDE can take with respect to a particular choice of topology. The first work studying the support of a stochastic process was [SV72] where the law of an SDE is characterised in terms of the supremum norm and the authors goal was to establish a Strong Maximum principle for a class of Elliptic Partial Differential Equations. This was later extended to a wide class of processes in [GP90]. Later, a support theorem with respect to the Hölder norm was established in [BAGL94], and for a much wider class of norms in [GNSS95]. These works laid the groundwork for the later publication [LQZ02] which studies the support of the solution law of a rough differential equation driven by a Gaussian white noise. In [FLS06] it is shown that the continuity of the Itô-Lyons map means that the proof of a support theorem can be reduced to establishing a characterisation of the support of the driving noise in an adequately rich topology.

Support theorem results have been key in some other applications, for example, a support theorem for SDEs with jump noise was crucial in showing *Exponential Ergodicity* in [Kul09]. One of the conditions the authors require is *Topological Irreducibility*, that for any two points, there is a path of the jump process that passes between them in finite time. This can be verified by finding an expression for the support of the law. Support theorems are also central in the establishment of *Stochastic Invariance principle*. A stochastic process is said to be invariant of a closed set $\mathcal{D} \subseteq \mathbb{R}^d$ if the solution starts and remains on the set $\mathcal{D}$ $\mathbb{P}$-almost surely $\forall t \in [0, T]$. This problem was first studied in [ADP90]. More recently, Stochastic Invariance has been studied in [Zab00], [BQRT10] and [FTT14]. In general, support theorems continue to draw attention from a wide range of academics, see [CF18], [CK19] and [HS19]. Lastly, a motivation to study support theorem results for McKean-Vlasov Equations is the recent link between this class of equations, deep learning (or *rich learning*) and ergodicity, see [HKR19].

4.1 Classical Support Theorm

The following useful method for proving a classical support theorem can be found in [MSS94].

Theorem 4.1.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space containing a Brownian motion and let E be a separable Banach space. Let $\mathcal{H}$ be the reproducing kernel Hilbert space of Brownian motion. Let $X : \Omega \to E$ be a random variable and let $\Phi : \mathcal{H} \to E$ be a measurable map.*

1. *Suppose there exists a sequence of random variables $H_n : \Omega \to \mathcal{H}$ such that for any $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}\Big[\|X(\cdot) - \Phi(H_n(\cdot))\|_E > \varepsilon\Big] = 0. \tag{4.1.1}$$

   Then $\operatorname{supp}(\mathcal{L}^X) \subset \overline{\Phi(\mathcal{H})}^E$.

2. *Suppose there exists a sequence of measure transforms T_n^h such that $\mathbb{P} \circ T_n^h$ is absolutely*

continuous with respect to $\mathbb{P}$ *and for any* $\varepsilon > 0$

$$\limsup_{n\to\infty} \mathbb{P}\Big[\|X(T_n^h(\cdot)) - \Phi(h)\|_E < \varepsilon\Big] > 0. \tag{4.1.2}$$

Then $\overline{\Phi(\mathcal{H})}^E \subset \mathrm{supp}(\mathcal{L}^X)$.

If both (4.1.1) *and* (4.1.2) *are satisfied, then* $\overline{\Phi(\mathcal{H})}^E = \mathrm{supp}(\mathcal{L}^X)$ *and* Φ *is called the* skeleton process *of the random variable* X*, see* [CFN97].

Equation (4.1.1) is sometimes referred to as the Wong Zakai implication due to its similarity with the Wong Zakai theorem. Equation (4.1.2) is sometimes referred to as the Cameron Martin implication because the proof involves exploiting the absolute continuity of Cameron Martin transforms on Wiener space.

4.2 Our Contributions

Proving a support theorem for McKean-Vlasov Equations is more challenging than verifying Equations (4.1.1) and (4.1.2). The knowledgeable reader will realise that for McKean-Vlasov Equations, the skeleton process is itself dependent on the law of the solution of the McKean-Vlasov Equation so the law must be known exogenously in order to solve any skeleton process path.

This is in contrast to the skeleton process used in [dRST19] where the measure dependency is replaced by a Dirac following the skeleton process driven by a constant 0 noise.

To *prove the support of McKean-Vlasov Equation* we develop a novel method by considering the sequences of pairs $(H_n, \mathcal{L}_n)_{n\in\mathbb{N}}$ and $(T_n^h, \mathcal{L}_n)_{n\in\mathbb{N}}$ where $(\mathcal{L}_n)_{n\in\mathbb{N}}$ is a sequence of measures that converge to the law of the McKean-Vlasov Equation. However, for each $n \in \mathbb{N}$, the skeleton process $\Phi(h, \mathcal{L}_n)$ driven by $\mathcal{L}_n$ and a reproducing kernel Hilbert space path h are not necessarily contained in the support even though they are a good approximation of a path that is contained in the support. Thus our statement for the support takes the form (see Theorem 14.2.4 below)

$$\mathrm{supp}(\mathcal{L}) = \bigcap_{n\in\mathbb{N}} \overline{\bigcup_{m\geq n} \Big\{\Phi(h, \mathcal{L}_m) : h \in \mathcal{H}\Big\}}^{\alpha-\text{Höl}}.$$

By solving the system of interacting rough differential equations driven by a Hölder quantization of the Brownian motion and exploiting the continuity properties of rough differential equations, we obtain a *deterministic, finite support measure* that approximates the law of the McKean-Vlasov Equation without having to solve the law explicitly. One could equivalently obtain the solution law by solving the non-linear Fokker Planck equation, but a novelty of this work is the obtention of the law without having to resort to PDE methods. We initiate our study by developing our results entirely within the framework established in [CL15].

The key advantage of this deterministic construction over the use of Empirical distribution used in McKean-Vlasov numerics is that we avoid all difficulties with characterising the support (a deterministic set) from random samples. For instance, the almost-sure rate of convergence for an Empirical distribution may, for a particular sample, be too poor to be of any effective use.

It is also worth emphasising that the rate of convergence that we obtain in Theorem 12.2.8 is, at face value, much slower than other well known methods for sampling a measure. The reason for this is we approximate in pathspace rather than for any fixed choice of time. Thus our quantization encodes both information about the path of a Brownian motion and the Hölder regularity.

In this thesis, we prove two support theorems, see Theorem 14.2.4 and Theorem 14.3.6. The first is for McKean-Vlasov Equations where the initial condition is deterministic while the second is a extension of this result for McKean-Vlasov Equations with random initial condition. The proof of the extension is simple and follows from [CFN97] so we focus predominantly on the first case.

Chapter 5

Small Ball Probabilities

Small Ball Probabilites, sometimes referred to as Small Deviations Principles, study the asymptotic behavour of the measure of a ball of radius $\varepsilon \to 0$. Given a measure $\mathcal{L}$ on a metric space (E, d) with Borel σ-algebra $\mathcal{B}$, we refer to the Small Ball Probability around a point x_0 as

$$\log\left(\mathcal{L}\left[\left\{x \in E : d(x, x_0) < \varepsilon\right\}\right]\right) \qquad \varepsilon \to 0.$$

This is in contrast to a Large Deviations Principle which considers the asymptotic behaviour for the quantity

$$\log\left(\mathcal{L}\left[\left\{x \in E : d(x, x_0) > a\right\}\right]\right) \qquad a \to \infty.$$

Large Deviations Principles have proved to be a powerful tool for quantifying the tails of Gaussian probability distributions that have been sucessfully explored and documented in recent years, see for example [Bog98, LT13] and references therein. Similar results have been extended to a wide class of probability distributions, see for example [Var84, DZ10]. However, the complexity of Small Ball Probabilities has meant there has been a generally slower growth in the literature. This is not to detract from their usefulness: there are many insightful and practical applications of Small Ball Probabilities to known problems, in particular the study of Compact operators, computation of Hausdorff dimension and the rate of convergence of Empirical and Quantized distributions.

As a motivational example, let $\mathcal{L}$ be a Gaussian measure on $\mathbb{R}^d$ with mean 0 and identity covariance matrix. Then

$$\mathcal{L}\left[\left\{x \in \mathbb{R}^d : |x|_2 < \varepsilon\right\}\right] = \frac{\Gamma(d/2) - \Gamma(d/2, \frac{\varepsilon^2}{2})}{\Gamma(d/2)} \sim \frac{2 \cdot \varepsilon^d}{\Gamma(d+1) \cdot 2^{d/2}} \qquad \varepsilon \to 0.$$

Therefore an application of l'Hôpital's rule yields

$$\mathfrak{B}_{0,2}(\varepsilon) = -\log\left(\mathcal{L}\left[\left\{x \in \mathbb{R}^d : |x|_2 < \varepsilon\right\}\right]\right) \sim d \cdot \log(\varepsilon^{-1}) \qquad \varepsilon \to 0.$$

Alternatively, using a different norm we have

$$\mathcal{L}\left[\left\{x \in \mathbb{R}^d : |x|_\infty < \varepsilon\right\}\right] = \mathrm{erf}\left(\tfrac{\varepsilon}{\sqrt{2}}\right)^d \sim \varepsilon^d \left(\tfrac{2}{\pi}\right)^{d/2} \qquad \varepsilon \to 0$$

and we get

$$\mathfrak{B}_{0,\infty}(\varepsilon) = -\log\left(\mathcal{L}\left[\left\{x \in \mathbb{R}^d : |x|_\infty < \varepsilon\right\}\right]\right) \sim d \cdot \log(\varepsilon^{-1}) \qquad \varepsilon \to 0.$$

5.1 Previous results for Small Ball Probabilities

Small ball probabilities encode the shape of the cumulative distribution function for a norm around 0. For a self-contained introduction to the theory of small ball probabilities and Gaussian inequalities, see [LS01].

Small ball probabilities for a Brownian motion with respect to the Hölder norm were first studied in [BR92]. Using the Cielsielski representation of Brownian motion, the authors are able to exploit the orthogonality of the Schauder wavelets in the Reproducing Kernel Hilbert space to represent the probability as a product of probabilities of 1 dimensional normal random variables. Standard analytic estimations of the Gauss error function provide an upper and lower bound for the probability and an expression for the limit for the probability as $\varepsilon \to 0$. Thus

$$\begin{aligned}\log\Big(\mathbb{P}\Big[\|W\|_\alpha < \varepsilon\Big]\Big) &= \log\Big(\prod_{(p,m)\in\Delta} \mathbb{P}\Big[2^{p(\alpha-1/2)} W_{pm} < \varepsilon\Big]\Big)\\ &= \sum_{(p,m)\in\Delta} \log\Big(\operatorname{erf}\Big(\tfrac{\varepsilon 2^{p(1/2-\alpha)}}{\sqrt{2}}\Big)\Big) \approx \varepsilon^{\frac{2}{1-2\alpha}} \qquad \text{as } \varepsilon \to 0.\end{aligned}$$

Later, the same results were extended to a large class of Gaussian processes under different assumptions for the covariance and different choices of Banach space norms, see for example [KL93b, KLS95, Sto96] and others.

In [DM14], the author studies some small ball probabilities for Levy Area of Brownian motion by treating it as a time-changed Brownian motion. However, there are no works studying small ball probabilities for Gaussian rough paths.

However, small ball probabilities have been used study the integrability of rough paths. In [CHLT15], while studying properties of the densities of rough differential equations driven by Gaussian rough paths the authors need to establish integrabilitity properties of the modulus of Hölder roughness for the driving noise. These are solved by computing its small ball probabilities which correspond to the tail distribution of one over the modulus of Hölder roughness.

The Metric Entropy of a set is a way of measuring the "Compactness" of a compact set. For a neat introduction to the study of Entropy and some of its applications, see [CS90] and [ET96]. The link between Small Ball Probabilities for Gaussian measures on Banach spaces is explored in [KL93a] and later extended in [LL99] to encompass the truncation of Gaussian measures. The same ideas to link the small ball probabilities and metric entropies for Stable processes are studied in [LL04, Aur07, ALL09]. Small ball probability results for Integrated Brownian motion, see [GHT03], were used to compute the metric entropy of k-monotone functions in [Gao08]. The link between the entropy of the convex hull of a set and the associated Gaussian measure is explored in [Gao04, Kle13]. For a recent survey on Gaussian measures and metric entropy, see [KL17].

There is a natural link between the Metric Entropy of the unit ball of the Reproducing Kernel Hilbert space of a Gaussian measure and the Quantization problem. Using the Large Deviations of the Gaussian measure, one can easily find a ball (in the RKHS) with measure $1-\varepsilon$ where $0 < \varepsilon \ll 1$. Given the ε entropy of this set, the centres of the minimal cover represent a very reasonable "guess" for an optimal quantization since the Gaussian measure conditioned on the closure of this set is "close" to uniform. For more details, see [GLP03, DFMS03]. Sharp estimates for Kolmogorov numbers, an equivalent measure to Metric Entropy, are demonstrated in [LP04].

More recently, Small Ball Probabilities have been applied to Baysian inference and machine learning, see for example [vdVvZ07, vdVvZ11, AILVZ08] and [WSS01].

5.2 Gaussian Correlation Inequalities

A key step in the proof of many small ball probability results is the use of a correlation inequality to lower or upper bound a probability of the intersection of many sets by a product of the probabilities of each set. Thus a challenging probability computation can be simplified by optimising over the choice of correlation strategically.

The Gaussian Correlation Inequality states that for any two symmetric convex sets A and B in a separable Banach space and for any centred Gaussian measure $\mathcal{L}$ on E,

$$\mathcal{L}[A \cap B] \geq \mathcal{L}[A]\mathcal{L}[B].$$

The first work which considers a special case of this result was conjectured in [DS55], while the first formal statement was made in [GEO$^+$72].

While the inequality remained unproven until recently, prominent works proving special examples and weaker statements include [Kha67, Šid68] (who independently proved the so called Šidák's Lemma), [Pit77] and [Li99]. The conjecture was proved in 2014 by Thomas Royen in a relatively uncirculated ArXiv submission [Roy14] and did not come to wider scientific attention for another three years in [LaM17].

Put simply, the idea is to minimise a probability for a collection of normally distributed random variables by varying the correlation. Applications of these inequalities are wide ranging and vital to the theory of Baysian inference.

5.3 Our Contributions

In order to extend the theory of Gaussian measures on Banach spaces to the framework of rough paths, we need to rephrase several well known Gaussian inequalities and prove several new correlation inequalites. This is done in Chapter 15. While technical, these results are stronger than we require and represent an extension of the theory of Correlation Inequalities to elements of the Wiener Itô chaos expansion.

The main contribution of Part V is the computation of Small Ball Probabilites for Gaussian rough paths with the rough path Hölder metric. These results are solved in Chapter 16. We remark that the discretisation of the Hölder norm in Lemma 16.1.1 was unknown to the author and may be of independent interest for future works on rough paths.

Finally, Chapter 17 demonstrates some applications of Theorem 16.0.1 following known methods that are adapted to the rough path setting. Of particular interest are Theorems 17.2.1 and 17.2.3 which provide an upper and lower bound for the rate of convergence for an empirical rough Gaussian measure.

Part II

Small Noise Limits of McKean-Vlasov Equations

Chapter 6

McKean-Vlasov equations with locally Lipschiz coefficients

The results of this Chapter can be found published in [dRST19, Section 3]. The first Lemma is a simple computation which we evaluate for the benefit of the reader.

Lemma 6.0.1. *Take δ_0. Then for any $\mu \in \mathcal{P}_2(E)$ we have $\mathbb{W}^{(2)}(\mu, \delta_0) = \big(\int_E y^2 \mu(dy) \big)^{1/2}$.*

Proof. Consider a random variable with law δ_0. We have $X : \Omega \to E$ with $\mathbb{P}[X \in A] = \delta_0(A)$ for any $A \subset E$. The σ-algebra generated by X is just $\{\Omega, \varnothing\}$. Let $\mu \in \mathcal{P}_2[E]$ be the law of a random variable $Y : \Omega \to E$ which generates a σ-algebra that X will be measurable with respect to. For any $B \in \sigma(Y)$, we have that $\mathbb{P}[\Omega \cap B] = \mathbb{P}[B] = 1\mathbb{P}[B] = \mathbb{P}[\Omega]\mathbb{P}[B]$ and $\mathbb{P}[B \cap \varnothing] = \mathbb{P}[\varnothing] = 0 = \mathbb{P}[B]\mathbb{P}[\varnothing]$. Hence $\sigma(X)$ and $\sigma(Y)$ are independent. Therefore we have that the joint density function of X and Y is just $\mu(dy)\delta_0(dx)$ and the conclusion follows. $\square$

6.1 Existence and Uniqueness

We start by extended previous results for Existence and Uniqueness of solutions of McKean-Vlasov Equations to the locally Lipschitz case. We work with general monotonicity assumptions without imposing coercivity restrictions. We also sharpen the integrability assumptions and leave it to the reader to verify that the proof in [Car16] can be sharpened.

Assumption 6.1.1. *Let $p \geq 2$. The progressively measurable maps $b : [0,T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d \times d'}$ satisfy that $\exists L > 0$ such that:*

1. *$Y_0 \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$ be independent of the Brownian motion.*
2. *Integrability: b and σ satisfy*
$$\mathbb{E}\Big[\Big(\int_0^T |b(t,0,\delta_0)| dt\Big)^p\Big], \mathbb{E}\Big[\int_0^T \Big|\sigma(t,0,\delta_0)\Big|^2 dt\Big)^{\frac{p}{2}}\Big] < \infty,$$
3. *σ is Lipschitz: $\forall t \in [0,T]$, $\forall x, x' \in \mathbb{R}^d$ and $\forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have*
$$|\sigma(t,x,\mu) - \sigma(t,x',\mu')| \leq L\Big(|x-x'| + \mathbb{W}^{(2)}(\mu,\mu')\Big),$$
4. *b satisfies the monotone growth condition in x and is Lipschitz in μ: $\forall t \in [0,T]$, $\forall x, x' \in \mathbb{R}^d$ and $\forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have that*
$$\langle x - x', b(t,x,\mu) - b(t,x',\mu) \rangle_{\mathbb{R}^d} \leq L|x-y|^2$$
$$\text{and} \quad |b(t,x,\mu) - b(t,x,\mu')| \leq L\mathbb{W}^{(2)}(\mu,\mu'),$$

5. *b is Locally Lipschitz with Polynomial Growth in x: $\exists q \in \mathbb{N}$ such that $q > 1$ and $\forall t \in [0,T]$, $\forall \mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\forall x, x' \in \mathbb{R}^d$ we have*

$$|b(t,x,\mu) - b(t,x',\mu)| \leq L(1 + |x|^{q-1} + |x'|^{q-1})|x - x'|.$$

Theorem 6.1.2. *Let $p \geq 2$. Suppose the drift and diffusion coefficients b, σ and initial distribution Y_0 satisfy Assumption 6.1.1. Then there exists a unique solution $Y \in \mathcal{S}^p$ to* (1.2.1) *and $\exists C > 0$ such that*

$$\mathbb{E}\Big[\|Y\|_\infty^p\Big] \leq C\left(\mathbb{E}\Big[|Y_0|^p\Big] + \mathbb{E}\Big[\Big(\int_0^T |b(t,0,\delta_0)|dt\Big)^p\Big] + \mathbb{E}\Big[\Big(\int_0^T |\sigma(s,0,\delta_0)|^2 ds\Big)^{\frac{p}{2}}\Big]\right)e^{CT}.$$

Proof. Consider the operator

$$\Xi : \mathcal{P}_2(C([0,T],\mathbb{R}^d)) \to \mathcal{P}_2(C([0,T],\mathbb{R}^d)),$$

where $\Xi(\mu) = \mathcal{L}^{Y^\mu}$ denotes the law of the SDE's solution Y^μ with dynamics

$$dY_t^\mu = b(t, Y_t^\mu, \mu_t)dt + \sigma(t, Y_t^\mu, \mu_t)dW_t, \quad Y_0^\mu = Y_0.$$

We start by showing that given some μ, a solution to the above SDE exists. Let $\mu \in \mathcal{P}_2(C([0,T],\mathbb{R}^d))$. Define

$$\hat{b}^\mu(t,x) = b(t,x,\mu_t), \quad \hat{\sigma}^\mu(t,x) = \sigma(t,x,\mu_t).$$

Then we have

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T |\hat{b}^\mu(t,0)|dt\Big)^p\Big] \\
&\quad \leq \mathbb{E}\Big[\Big(\int_0^T |b(t,0,\delta_0)| + L \cdot \mathbb{W}^{(2)}(\mu_t,\delta_0)dt\Big)^p\Big] \\
&\quad \leq 2^{p-1}\mathbb{E}\Big[\Big(\int_0^T |b(t,0,\delta_0)|dt\Big)^p\Big] + 2^{p-1}L^pT^p \cdot \mathbb{W}^{(2)}(\mu,\delta_0)^p < \infty,
\end{aligned}$$

and similarly

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T |\hat{\sigma}^\mu(t,0)|^2 dt\Big)^{\frac{p}{2}}\Big] \\
&\quad \leq 2^{p-1}\mathbb{E}\Big[\Big(\int_0^T |\sigma(t,0,\delta_0)|^2 dt\Big)^{\frac{p}{2}}\Big] + 2^{p-1}L^pT^p \cdot W^{(2)}(\mu,\delta_0)^p < \infty.
\end{aligned}$$

Also we have that $\hat{b}^\mu(t,x)$ is locally Lipschitz, satisfies a monotone growth condition and $\hat{\sigma}^\mu(t,x)$ has Lipschitz growth in its spacial variables. Therefore, by the methods in [Mao08, Theorem 3.6], we have that a unique solution exists in $\mathcal{S}^p([0,T])$. Since $p \geq 2$, we can conclude that $\mathcal{L}^{Y^\mu} \in \mathcal{P}_2(C([0,T],\mathbb{R}^d))$.

Using Itô's formula, we have that

$$\begin{aligned}
\mathbb{W}^{(2)}(\Xi(\mu),\Xi(\nu))^2 &\leq \mathbb{E}\Big[\sup_{t\in[0,T]} |Y_t^\mu - Y_t^\nu|^2\Big] \\
&\leq 2\mathbb{E}\Big[\int_0^T |\langle Y_s^\mu - Y_s^\nu, b(s,Y_s^\mu,\mu_s) - b(s,Y_s^\nu,\nu_s)\rangle| ds\Big] && (6.1.1) \\
&\quad + 2\mathbb{E}\Big[\sup_{t\in[0,T]} \int_0^t \Big\langle Y_s^\mu - Y_s^\nu, [\sigma(s,Y_s^\mu,\mu_s) - \sigma(s,Y_s^\nu,\nu_s)]dW_s\Big\rangle\Big] && (6.1.2) \\
&\quad + \mathbb{E}\Big[\int_0^T \Big|\sigma(s,Y_s^\mu,\mu_s) - \sigma(s,Y_s^\nu,\nu_s)\Big|^2 ds\Big]. && (6.1.3)
\end{aligned}$$

Firstly, we apply the monotonicity and Lipschitz properties to get

$$\begin{aligned}(6.1.1) \leq & 2L\mathbb{E}\Big[\int_0^T |Y_s^\mu - Y_s^\nu|^2 ds\Big] + 2L\mathbb{E}\Big[\int_0^T |Y_s^\mu - Y_s^\nu| \cdot \mathbb{W}^{(2)}(\mu_s, \nu_s) ds\Big] \\ \leq & 2L\int_0^T \mathbb{E}\Big[\|Y^\mu - Y^\nu\|_{\infty,s}\Big] ds + \frac{\mathbb{E}\Big[\|Y^\mu - Y^\nu\|_\infty^2\Big]}{3} + 3L^2 \int_0^T \mathbb{W}^{(2)}(\mu_s, \nu_s)^2 ds.\end{aligned}$$

Secondly, we use the Burkholder Davis Gundy inequality and Lipschitz properties

$$\begin{aligned}(6.1.2) \leq & 2\mathbb{E}\Big[\Big(\int_0^T \Big|(Y_s^\mu - Y_s^\nu)^T \Big(\sigma(s, Y_s^\mu, \mu_s) - \sigma(s, Y_s^\nu, \nu_s)\Big)\Big|^2 ds\Big)^{\frac{1}{2}}\Big] \\ \leq & 2\mathbb{E}\Big[\|Y^\mu - Y^\nu\|_\infty \Big(\int_0^T |\sigma(s, Y_s^\mu, \mu_s) - \sigma(s, Y_s^\nu, \nu_s)|^2 ds\Big)^{\frac{1}{2}}\Big] \\ \leq & \frac{\mathbb{E}\Big[\|Y^\mu - Y^\nu\|_\infty^2\Big]}{3} + 6L^2 \int_0^T \mathbb{E}\Big[\|Y^\mu - Y^\nu\|_{\infty,s}^2\Big] ds + 6L^2 \int_0^T \mathbb{W}^{(2)}(\mu_s, \nu_s)^2 ds.\end{aligned}$$

Thirdly, using the Lipschitz properties again we get

$$(6.1.2) \leq 2L^2 \int_0^T \mathbb{E}\Big[\|Y^\mu - Y^\nu\|_{\infty,s}^2\Big] ds + 2L^2 \int_0^T \mathbb{W}^{(2)}(\mu_s, \nu_s)^2 ds.$$

Combining (6.1.1), (6.1.2) and (6.1.3) gives that

$$\frac{\mathbb{E}\Big[\|Y^\mu - Y^\nu\|_\infty^2\Big]}{3} \leq (8L^2 + 2L) \int_0^T \mathbb{E}\Big[\|Y^\mu - Y^\nu\|_{\infty,s}^2\Big] ds + 11L^2 \int_0^T \mathbb{W}^{(2)}(\mu_s, \nu_s)^2 ds.$$

Applying Grönwall to this yields a control to the initial inequality

$$\mathbb{W}^{(2)}(\Xi(\mu), \Xi(\nu))^2 \leq \mathbb{E}\Big[\|Y^\mu - Y^\nu\|_\infty^2\Big] \leq K \int_0^T \mathbb{W}_s^{(2)}(\mu, \nu)^2 ds,$$

where $K = 11L^2 e^{(24L^2+6L)T}$. Applying Ξ inductively j times yields

$$\begin{aligned}\mathbb{W}^{(2)}\Big(\Xi^j(\mu), \Xi^j(\nu)\Big)^2 \leq & K^j \int_0^T \int_0^{t_1} \dots \int_0^{t_{j-1}} \mathbb{W}_{t_j}^{(2)}(\mu, \nu)^2 dt_1 \dots dt_j \\ \leq & K^j \int_0^T \frac{(T - t_j)^{j-1}}{(j-1)!} \mathbb{W}_{t_j}^{(2)}(\mu, \nu)^2 dt_j \leq \frac{K^j T^j}{j!} \mathbb{W}^{(2)}(\mu, \nu)^2.\end{aligned}$$

Choosing j large enough ensures that Ξ^j is a contraction operator. Therefore, Ξ has a unique fixed point. Hence we conclude that the Picard sequence of random processes $Y_t^0 = Y_0$ and

$$dY_t^n = b\big(t, Y_t^n, \mathcal{L}_t^{Y^{n-1}}\big) dt + \sigma\big(t, Y_t^n, \mathcal{L}_t^{Y^{n-1}}\big) dW_t,$$

converges in $\mathcal{S}^2$ and the limit solves the McKean-Vlasov Equation (1.2.1). From this we conclude that a unique solution exists.

Step 2: Moment calculations. Recall the dynamics of Y from (1.2.1). By Itô's formula we have

$$\begin{aligned}|Y_t|^p = & |Y_0|^p + p\int_0^t |Y_s|^{p-2} \langle Y_s, b(s, Y_s, \mathcal{L}_s^Y)\rangle ds + p\int_0^t |Y_s|^{p-2} \langle Y_s, \sigma(s, Y_s, \mathcal{L}_s^Y) dW_s\rangle \\ & + \frac{p}{2}\int_0^t |Y_s|^{p-2} \Big|\sigma(s, Y_s, \mathcal{L}_s^Y)\Big|^2 ds + \frac{p(p-2)}{2}\int_0^t |Y_s|^{p-4} \cdot \Big|Y_s^T \sigma(s, Y_s, \mathcal{L}_s^Y)\Big|^2 ds.\end{aligned}$$

Therefore

$$\mathbb{E}\Big[\|Y\|_\infty^p\Big] \leq \mathbb{E}\Big[|Y_0|^p\Big] + p\mathbb{E}\Big[\int_0^T |Y_s|^{p-2}\big|\langle Y_s, b(s, Y_s, \mathcal{L}_s^Y)\rangle\big| ds\Big] \tag{6.1.4}$$

$$+ p\mathbb{E}\Big[\sup_{0\leq t\leq T}\int_0^t |Y_s|^{p-2}\langle Y_s, \sigma(s, Y_s, \mathcal{L}_s^Y)dW_s\rangle\Big] \tag{6.1.5}$$

$$+ \frac{p}{2}\mathbb{E}\Big[\int_0^T |Y_s|^{p-2}\Big|\sigma(s, Y_s, \mathcal{L}_s^Y)\Big|^2 ds\Big] \tag{6.1.6}$$

$$+ \frac{p(p-2)}{2}\mathbb{E}\Big[\int_0^T |Y_s|^{p-4}\cdot\Big|Y_s^T\sigma(s, Y_s, \mathcal{L}_s^Y)\Big|^2 ds\Big]. \tag{6.1.7}$$

By the triangle property we have

$$(6.1.4) \leq p\mathbb{E}\Big[\int_0^T |Y_s|^{p-2}\langle Y_s, b(s, Y_s, \mathcal{L}_s^Y) - b(s, 0, \mathcal{L}_s^Y)\rangle ds\Big] \tag{6.1.8}$$

$$+ p\mathbb{E}\Big[\int_0^T |Y_s|^{p-2}\langle Y_s, b(s, 0, \mathcal{L}_s^Y) - b(s, 0, \delta_0)\rangle ds\Big] \tag{6.1.9}$$

$$+ p\mathbb{E}\Big[\int_0^T |Y_s|^{p-2}\langle Y_s, b(s, 0, \delta_0)\rangle ds\Big]. \tag{6.1.10}$$

Using the monotone property of b yields (6.1.8) $\leq pL\int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^p\Big]ds$. Using the Lipschitz property of b in the distribution variable and Lemma 6.0.1 yields

$$(6.1.9) \leq pL\int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^{p-1}\Big]\mathbb{E}\Big[\|Y\|_{\infty,s}^2\Big]^{\frac{1}{2}} ds \leq pL\int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^p\Big]ds.$$

Using the integrability properties of b yields

$$\begin{aligned}(6.1.10) \leq& \mathbb{E}\Big[\|Y\|_\infty^{p-1}\int_0^T |b(s, 0, \delta_0)|ds\Big]\\ \leq& \frac{\mathbb{E}\Big[\|Y\|_\infty^p\Big]}{n} + n^{p-1}(p-1)^{p-1}\mathbb{E}\Big[\Big(\int_0^T |b(s, 0, \delta)|ds\Big)^p\Big]\end{aligned}$$

where $n \in \mathbb{N}$ which will be chosen later.

By the Burkholder-Davis-Gundy inequality, the Lipschitz properties and Lemma 6.0.1 we have

$$\begin{aligned}(6.1.5) \leq& pC_1\mathbb{E}\Big[\Big(\int_0^T |Y_s|^{2p-4}\Big|Y_s^T\sigma(s, Y_s, \mathcal{L}_s^Y)\Big|^2 ds\Big)^{\frac{1}{2}}\Big]\\ \leq& \frac{\mathbb{E}\Big[\|Y\|_{\infty,s}^p\Big]}{n} + 3p^2C_1^2nL^2\Big(\int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^p\Big]ds + \int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^{p-2}\Big]\cdot\mathbb{E}\Big[\|Y\|_{\infty,s}^2\Big]ds\Big)\end{aligned} \tag{6.1.11}$$

$$+ 3p^2C_1^2n\mathbb{E}\Big[\int_0^T |Y_s|^{p-2}\Big|\sigma(s, 0, \delta_0)\Big|^2 ds\Big]. \tag{6.1.12}$$

Terms (6.1.11) are dealt with in the same way as terms (6.1.8) and (6.1.9). For (6.1.12) we proceed as follows

$$\begin{aligned}(6.1.12) \leq& \mathbb{E}\Big[\|Y\|_\infty^{p-2}\Big(3p^2C_1^2n\int_0^T \Big|\sigma(s, 0, \delta_0)\Big|^2 ds\Big)\Big]\\ \leq& \frac{\mathbb{E}\Big[\|Y\|_\infty^p\Big]}{n} + 2\cdot 3^{\frac{p}{2}}\cdot n^{p-1}C_1^p(p-2)^{\frac{p-2}{2}}p^{\frac{p}{2}}\mathbb{E}\Big[\Big(\int_0^T \Big|\sigma(s, 0, \delta_0)\Big|^2 ds\Big)^{\frac{p}{2}}\Big].\end{aligned}$$

Thirdly, we have

$$
\begin{aligned}
(6.1.6)+(6.1.7) \leq & \frac{p(p-1)}{2}\mathbb{E}\Big[\int_0^T |Y_s|^{p-2}\Big|\sigma(s,Y_s,\mathcal{L}_s^Y)\Big|^2 ds\Big] \\
\leq & \frac{3p(p-1)L^2}{2}\Big(\int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^p\Big]ds + \int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^{p-2}\Big]\cdot\mathbb{E}\Big[\|Y\|_{\infty,s}^2\Big]ds\Big) \\
& + \frac{3p(p-1)}{2}\mathbb{E}\Big[\|Y\|_\infty^{p-2}\int_0^T \Big|\sigma(s,0,\delta_0)\Big|^2 ds\Big] \qquad (6.1.13)
\end{aligned}
$$

and

$$
(6.1.13) \leq \frac{\mathbb{E}\Big[\|Y\|_\infty^p\Big]}{n} + \Big(\frac{n(p-2)}{2}\Big)^{\frac{p-2}{2}} \cdot \Big(3(p-1)\Big)^{\frac{p}{2}}\mathbb{E}\Big[\Big(\int_0^T \Big|\sigma(s,0,\delta_0)\Big|^2 ds\Big)^{\frac{p}{2}}\Big].
$$

Hence we choose $n=5$ and this can all be rearranged to get

$$
\begin{aligned}
\frac{\mathbb{E}\Big[\|Y\|_\infty^p\Big]}{5} \leq & \mathbb{E}\Big[|Y_0|^p\Big] + \tilde{C}_1\mathbb{E}\Big[\Big(\int_0^T \Big|\sigma(s,0,\delta_0)\Big|^2 ds\Big)^{\frac{p}{2}}\Big] \\
& + \tilde{C}_2\mathbb{E}\Big[\Big(\int_0^T |b(s,0,\delta_0)|ds\Big)^p\Big] + \tilde{C}_3\int_0^T \mathbb{E}\Big[\|Y\|_{\infty,s}^p\Big]ds,
\end{aligned}
$$

where the constants $\tilde{C}_1$, $\tilde{C}_2$ and $\tilde{C}_3$ are dependent only on p and L. Applying Grönwall's lemma provides us with the p moment upper bound

$$
\mathbb{E}\Big[\|Y\|_\infty^p\Big] \leq 5\Bigg(\mathbb{E}\Big[|Y_0|^p\Big] + \tilde{C}_1\mathbb{E}\Big[\Big(\int_0^T \Big|\sigma(s,0,\delta_0)\Big|^2 ds\Big)^{\frac{p}{2}}\Big] + \tilde{C}_2\mathbb{E}\Big[\Big(\int_0^T |b(s,0,\delta_0)|ds\Big)^p\Big]\Bigg)e^{\tilde{C}_3 T}.
$$

□

6.2 Continuity in time behavior

We next give results describing time-continuity for the process and its law in the appropriate topologies. We use the standard notation that $Y_{s,t} = Y_t - Y_s$.

Proposition 6.2.1. *Let Y be the solution of* (1.2.1) *satisfying Assumption 6.1.1 where $q \in \mathbb{N}$ is the order of the polynomial growth of b. Let $n \in \mathbb{N}$ and $n \geq 2$ and additionally assume that*

$$
\mathbb{E}\Big[\sup_{t\in[0,T]} |b(t,0,\delta_0)|^{nq}\Big],\ \mathbb{E}\Big[\sup_{t\in[0,T]} \Big|\sigma(t,0,\delta_0)\Big|^{\frac{nq}{2}}\Big] < \infty.
$$

Then for every $t,s \in [0,T]$

$$
\mathbb{W}^{(n)}\Big(\mathcal{L}_t^Y, \mathcal{L}_s^Y\Big) \leq \mathbb{E}\Big[\Big|Y_{s,t}\Big|^n\Big]^{\frac{1}{n}} \lesssim |t-s|^{\frac{1}{2}}.
$$

Proof. The proposition's conditions mean that by the arguments in the proof of Theorem 6.1.2 we have $\mathbb{E}[\|Y\|_\infty^{nq}] < \infty$. Take $0 \leq s \leq t \leq T < \infty$ and a natural number $n \geq 2$. We have

$$
|Y_{s,t}|^n \leq \Big|\int_s^t b(r,Y_r,\mathcal{L}_r^Y)dr + \int_s^t \sigma(r,Y_r,\mathcal{L}_r^Y)dW_r\Big|^n.
$$

We use the growth condition of b and the Lipschitz property of σ and apply the Minkowski

Inequality to get

$$
\begin{aligned}
&\mathbb{E}\Big[|Y_{s,t}|^n\Big]^{\frac{1}{n}} \\
&\quad\leq \mathbb{E}\Big[\Big|\int_s^t b(r,Y_r,\mathcal{L}_r^Y)dr\Big|^n\Big]^{\frac{1}{n}} + \mathbb{E}\Big[\Big|\int_s^t \sigma(r,Y_r,\mathcal{L}_r^Y)dW_r\Big|^n\Big]^{\frac{1}{n}} \\
&\quad\leq \mathbb{E}\Big[\Big|\int_s^t |b(r,Y_r,\mathcal{L}_r^Y)dr\Big|^n\Big]^{\frac{1}{n}} + \mathbb{E}\Big[\Big|\int_s^t |\sigma(r,Y_r,\mathcal{L}_r^Y)|^2 dr\Big|^{\frac{n}{2}}\Big]^{\frac{1}{n}} \\
&\quad\leq |t-s|\mathbb{E}\Big[\Big(\|b(\cdot,0,\delta_0)\|_\infty + L\|Y\|_\infty^q + \mathbb{E}\Big[\|Y\|^2\Big]^{\frac{1}{2}}\Big)^n\Big]^{\frac{1}{n}} \\
&\qquad + |t-s|^{\frac{1}{2}}\mathbb{E}\Big[\Big(\|\sigma(\cdot,0,\delta_0)\|_\infty + L\|Y\|_\infty + \mathbb{E}\Big[\|Y\|_\infty^2\Big]^{\frac{1}{2}}\Big)^n\Big]^{\frac{1}{n}} \lesssim |t-s|^{\frac{1}{2}},
\end{aligned}
$$

From the 1st part of the proposition, we have $\mathbb{E}\big[\,|Y_{s,t}|^{2p}\big] \lesssim |t-s|^p$. The results now follow by applying Kolmogorov's Continuity criterion in a standard fashion. □

Corollary 6.2.2. *Let Y be the solution of* (1.2.1) *under Assumption 6.1.1 and suppose additionally that $\forall n \in \mathbb{N}$ we have*

$$
\mathbb{E}\Big[\|b(\cdot,0,\delta_0)\|_\infty^n\Big] < \infty, \quad \mathbb{E}\Big[\|\sigma(\cdot,0,\delta_0)\|_\infty^n\Big] < \infty.
$$

Then there is a modification of Y, $\tilde{Y}$, which is sample-continuous, almost surely equal to Y and α-Hölder continuous for $\alpha < 1/2$.

Proof. Under these stronger conditions we have $\forall n \in \mathbb{N}$ that $\mathbb{E}[\,|Y_{s,t}|^n] \lesssim |t-s|^{n/2}$. Therefore, we apply the Kolmogorov Continuity Criterion from [Øks03, Theorem 2.2.3] and conclude. □

The final result concerns C^1-regularity (in time) of the expected value of maps of the McKean-Vlasov Equation.

Proposition 6.2.3 (Regularity in time). *Let $\phi \in C^{1,2}([0,T] \times \mathbb{R}^d)$ and suppose that ϕ, its first derivative, $\nabla_x\phi(\cdot,\cdot)$, and Hessian, $H[\phi](\cdot,\cdot)$, have polynomial growth such that for some $r > 0$ and some $K > 0$*

$$
\max\Big\{\big|\frac{\partial\phi}{\partial t}(t,y)\big|, \big|\nabla_x\phi(t,y)\big|, \big|H[\phi](t,y)\big|\Big\} \leq K\big(1+|y|^r\big).
$$

Suppose that Y is the solution to (1.2.1) *under Assumptions 6.1.1 with $p := \max\{r+q, 2r+2\}$ (q is the polynomial growth of b) and hence $Y \in \mathcal{S}^p$.*

Then $t \mapsto \mathbb{E}[\phi(t,Y(t))] \in C^1$ and

$$
\begin{aligned}
\partial_t\mathbb{E}[\phi(t,Y(t))] =& \mathbb{E}\Big[\frac{\partial\phi}{\partial t}(t,Y_t)\Big] + \mathbb{E}\Big[\nabla\phi(t,Y_t)^T \cdot b(t,Y_t,\mathcal{L}_t^Y)\Big] \\
&+ \mathbb{E}\Big[Tr\Big(\sigma(t,Y_t,\mathcal{L}_t^Y)^T \cdot H\Big[\phi\Big](t,Y_t)\cdot\sigma(t,Y_t,\mathcal{L}_t^Y)\Big)\Big].
\end{aligned}
$$

Proof. Use Itô's formula on $\phi(t,Y_t)$, integrate over $[0,t]$ and take expectations. By the integrability/growth assumptions on b and σ, we have $Y \in \mathcal{S}^p$ and in particular $Y \in \mathcal{S}^{2r+2}$. Combining with the polynomial growth of $\nabla\phi(\cdot,\cdot)$ in its spatial variable we easily conclude that the stochastic integral $\int_0^\cdot \nabla\phi(s,Y_s)\sigma(s,Y_s,\mathcal{L}_s^Y))dW_s$ is a square-integrable martingale and hence it vanishes under the expectation.

In the previous results we have shown continuity in time of Y and $\mathcal{L}^Y$ in the appropriate metrics. This, combined with the continuity of b and σ in their variables plus the integrability results, allows to apply Fubini and swap expectations and integrals. Lastly, using the continuity/integrability properties of the involved terms again (notice that here one requires $Y \in \mathcal{S}^{r+q}$), one can compute the time derivative of $t \mapsto \mathbb{E}[\phi(t,Y_t)]$ via the Leibniz differentiation rule for integrals. This yields the lemma's formula. □

Chapter 7

Large Deviations Principle in the Hölder Topology

In this chapter, we investigate the family of d-dimensional McKean-Vlasov Equations indexed by the parameter $\varepsilon > 0$,

$$X_t^\varepsilon = x + \int_0^t b_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})ds + \sqrt{\varepsilon}\int_0^t \sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})dW_s. \tag{7.0.1}$$

We derive a Freidlin-Wentzell Large Deviations Principle for equations of the form (7.0.1) with respect to the Hölder-norm. Throughout we make use of several known sources: [DZ10], [HIP08] and [BAL94]. Aside from the LDP results, contributions of this Chapter also include the techniques needed to deal directly with the law of the solution process inside the coefficients while avoiding measure arguments; time dependency of the coefficients is included.

The main body of work of this Chapter is proving Proposition 7.2.3 (see below). It is then a well known result of Large Deviations theory that this allows one to transfer an LDP result from a coarse topology to a finer one (see Theorem 7.2.2). For completeness, we additionally include a proof of Theorem 7.2.2.

The results of this Chapter can be found published in [dRST19, Section 4].

7.1 Auxilliary Results

In order to prove Proposition 7.2.3, we demonstrate some technical Lemmas that are inspired by the methods of [BAL94].

Lemma 7.1.1. *Let $(W_t)_{t\in[0,T]}$ be a d'-dimensional Brownian motion. There exists a constant $C' > 0$ which is independent of d' and α such that $\forall u > 0$ and $\forall K \in C([0,T];\mathbb{R}^{d\times d'})$ such that*

$$\mathbb{P}\bigg[\Big\| \int_0^\cdot K_s dW_s\Big\|_\alpha \geq u, \|K\|_\infty \leq 1\bigg] \leq C' \exp\Big(\frac{-u^2}{C'}\Big).$$

Proof of Lemma 7.1.1. Let $\|K\|_\infty \leq 1$. In the case where K is deterministic, the stochastic integral of K is clearly normally distributed and the result is clear. For K not deterministic, we do not know the distribution of the stochastic integral.

Using the equivalent definition of Hölder norms in terms of a Schauder expansion (see

Theorem A.1.2), we have that

$$\begin{aligned}\mathbb{P}\Big[\Big\|\int_0^{\cdot} K_s dW_s\Big\|_\alpha \geq u, \|K\|_\infty \leq 1\Big] &= \mathbb{P}\Big[\sup_{(p,m)\in\Lambda}\Big|\int_0^T H_{pm}(s)K_s dW_s\Big| \geq u, \|K\|_\infty \leq 1\Big] \\ &\leq \frac{\mathbb{E}\Big[\exp\Big(\lambda \sup_{(p,m)\in\Lambda}\Big|\int_0^T H_{pm}(s)K_s dW_s\Big|\Big)\mathbb{1}_{\{\|K\|_\infty<1\}}\Big]}{\exp(\lambda u)} \\ &\leq \sup_{(p,m)\in\Lambda}\frac{\mathbb{E}\Big[\exp\Big(\lambda\Big|\int_0^T H_{pm}(s)K_s dW_s\Big|\Big)\Big]}{\exp(\lambda u)}\end{aligned}$$

where the supremum can come outside the expectation by Beppo Levi Theorem since the random variables are all positive. Temporarily, consider the process $Y_t = \int_0^t H_{pm}(s)K_s dW_s$. Using Itô's formula we get that

$$\mathbb{E}[|Y_t|^n] = \int_0^t \mathbb{E}\Big[\frac{n(n-1)}{2}|Y_s|^{n-2}H_{pm}(s)^2K_s^2\Big]ds.$$

Y_t is a martingale, since H_{pm} and K are bounded, with $Y_0 = 0$ so $\mathbb{E}[Y_t] = 0$. By the Itô Isometry, the second moment of Y_t is

$$\mathbb{E}[Y_t^2] = \mathbb{E}\Big[\int_0^t H_{pm}(s)^2K_s^2 ds\Big] \leq \begin{cases} 0 & 0 \leq t \leq \frac{(m-1)T}{2^p} \\ \frac{2^p}{T}(t - \frac{(m-1)T}{2^p}) & \frac{(m-1)T}{2^p} < t < \frac{mT}{2^p} \\ 1 & \frac{mT}{2^p} \leq t \leq 1 \end{cases}$$

Therefore, by induction on n we see

$$\mathbb{E}[Y_t^{2n}] \leq \begin{cases} 0 & 0 \leq t \leq \frac{(m-1)T}{2^p} \\ \frac{(2n)!}{n!2^n}\Big(\frac{2^p}{T}\Big)^n(t - \frac{(m-1)T}{2^p})^n & \frac{(m-1)T}{2^p} < t < \frac{mT}{2^p} \\ \frac{(2n)!}{n!2^n} & \frac{mT}{2^p} \leq t \leq T \end{cases}.$$

For the odd moments of $|Y_t|$, we first use the Burkholder-Davies-Gundy Inequality to say

$$\begin{aligned}\mathbb{E}[|Y_t|] &\leq C_1\mathbb{E}\Big[\Big(\int_0^t H_{pm}(s)^2K_s^2 ds\Big)^{\frac{1}{2}}\Big] \\ &\leq \begin{cases} 0 & 0 \leq t \leq \frac{(m-1)T}{2^p} \\ C_1 2^{p/2}(t - \frac{(m-1)T}{2^p})^{1/2} & \frac{(m-1)T}{2^p} < t < \frac{mT}{2^p} \\ C_1 & \frac{mT}{2^p} \leq t \leq T \end{cases},\end{aligned}$$

and by induction on n again we see that

$$\mathbb{E}[|Y_t|^{2n+1}] \leq \begin{cases} 0 & 0 \leq t \leq \frac{(m-1)T}{2^p} \\ C_1 n!2^n(t - \frac{(m-1)T}{2^p})^{\frac{2n+1}{2}}2^{\frac{2n+1}{2}} & \frac{(m-1)T}{2^p} < t < \frac{mT}{2^p} \\ C_1 n!2^n & \frac{mT}{2^p} \leq t \leq T \end{cases}$$

Hence $\mathbb{E}[|Y_t|^{2n}] \leq (C_1 \vee 1)\frac{(2n)!}{n!2^n}$ and $\mathbb{E}[|Y_t|^{2n+1}] \leq (C_1 \vee 1)n!2^n$. The upper bounds for these moments are the same as the moments of a Half normal distribution with variance 1 up to a multiplicative constant. Therefore, we can upper bound the moment generating function of the random variable $|Y_1|$ using the moment generating function of a half normal random variable. If Z is half normally distributed with variance a, we have

$$\mathbb{E}\Big[\exp(\lambda Z)\Big] = \int_0^\infty \frac{2}{\sqrt{2\pi a^2}}e^{\lambda x}\exp\Big(\frac{-x^2}{2a^2}\Big)dx \leq 4\exp\Big(\frac{\lambda a^2}{2}\Big).$$

Therefore $\mathbb{E}\Big[\exp\Big(\lambda\Big|\int_0^1 H_{pm}(s)K_s dW_s\Big|\Big)\Big] \lesssim \exp\Big(\frac{\lambda^2}{2}\Big)$ and hence

$$\mathbb{P}\Big[\|\int_0^\cdot K_s dW_s\|_\alpha \geq u, \|K\|_\infty \leq 1\Big] \lesssim \exp\Big(\frac{\lambda^2}{2} - \lambda u\Big) \lesssim \exp\Big(\frac{-u^2}{2}\Big),$$

by choosing λ to minimize the equation since the choice of λ was arbitrary ($\lambda = u$). □

The following results are of their own independent interest and can be found in [BAL94, *Lemme* 1 p.196] with token proofs. We provide a full proof for the benefit of the reader.

Lemma 7.1.2 ([BAL94])**.** *Let $(W_t)_{t\in[0,T]}$ be a d'-dimensional Brownian motion. Then, there exists a constant $C > 0$ which is independent of m such that $\forall u, v > 0$*

$$\mathbb{P}\Big[\|W\|_\alpha \geq u, \|W\|_\infty \leq v\Big] \leq C \max\Big(1, \Big(\frac{u}{v}\Big)^{1/\alpha}\Big) \exp\Big(\frac{-1}{C}\frac{u^{1/\alpha}}{v^{(1/\alpha)-2}}\Big).$$

Proof of Lemma 7.1.2. Consider a Brownian motion W satisfying the constraint $\|W\|_\infty \leq v$. We use methods from [HIPP14] to represent the α-Hölder norm in terms of a supremum of Fourier coefficients generated by Schauder functions. By direct calculation, one can dominate the Fourier coefficients

$$|W_{pm}| = \sqrt{\tfrac{2^p}{T}}\Big|2W_{\frac{(2m-1)T}{2^{p+1}}} - W_{\frac{mT}{2^p}} - W_{\frac{(m-1)T}{2^p}}\Big| \leq \sqrt{\tfrac{2^p}{T}}4v.$$

If we also restrict that $\|W\|_\alpha = \sup_{p,m}|W_{pm}|2^{p(\alpha-1/2)} \geq u$ and search for values of p and m which do not yield a contradiction. Observe that we require $u \geq 4v2^{\alpha p}$. If we consider a p where this was not true, we would have that $W_{pm} < u$. The supremum of all W_{pm} is still be greater than u, but this value of p could be removed from the collection over which the supremum is taken over without affecting the measure of the event. Let p_0 be the least such relevant p, defined as $p_0 := \inf\{p \in \mathbb{N}; 2^{\alpha p} \geq u/(4v)\}$. Then for an arbitrary choice of $\lambda > 0$, we have

$$\begin{aligned}
&\mathbb{P}\Big[\|W\|_\alpha \geq u, \|W\|_\infty \leq v\Big] \\
&= \mathbb{P}\Big[\sup_{p\geq p_0, m} 2^{p(\alpha-1/2)}|W_{pm}| \geq u\Big] \leq \frac{\sup_{p\geq p_0}\mathbb{E}\Big[\exp(\lambda 2^{p(\alpha-1/2)}|W_{pm}|)\Big]}{\exp(\lambda u)} \\
&\leq \sup_{p\geq p_0} 2\exp\Big(\tfrac{\lambda^2 2^{p(2\alpha-1)}}{2} - \lambda u\Big) \leq 2\exp\Big(\tfrac{-u^2 2^{p_0(1-2\alpha)}}{2}\Big),
\end{aligned}$$

where for the last line we choose $\lambda = u2^{p(1-2\alpha)}$ to minimize the expression (since λ is arbitrary). From the definition of p_0 we have

$$2^{p_0(1-2\alpha)} \geq \Big(\frac{u}{4v}\Big)^{\frac{1}{\alpha}-2},$$

and substituting this in yields the final result. □

7.2 The main result

Recall the stochastic process (7.0.1). We introduce the so-called skeleton operator Φ for the McKean-Vlasov Equation (7.0.1) on the Cameron Martin Space $\mathcal{H}$, in other words, $\Phi : \mathcal{H} \to C([0,T];\mathbb{R}^d)$

$$\Phi(h)_t = x + \int_0^t b(s, \Phi(h)_s, \delta_{\Phi(0)_s})ds + \int_0^t \sigma(s, \Phi(h)_s, \delta_{\Phi(0)_s})\dot{h}_s ds. \tag{7.2.1}$$

The operator Φ for each $h \in \mathcal{H}$ outputs the unique solution to the above ODE.

Following the same method as in [dRST19, Lemma 4.3] and using the Hölder inequality, one can see that

$$|\Phi(h)_{s,t}| \leq O\Big(|t-s|\Big) + M|t-s|^{\frac{1}{2}}\|h\|_{\mathcal{H}} \leq O\Big(|t-s|^{\frac{1}{2}}\Big),$$

so $\Phi(h) \in C^{\frac{1}{2}}([0,T];\mathbb{R}^d)$. We are now able to state the two main results of this section:

Assumption 7.2.1. *Let $\varepsilon > 0$. Let $b, b_\varepsilon : [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$, $\sigma, \sigma_\varepsilon : [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d'}$ (deterministic maps) and $x \in \mathbb{R}^d$.*

As $\varepsilon \searrow 0$, let the maps b_ε converge uniformly to b and σ_ε converge uniformly to σ. Let b and σ satisfy Assumption 6.1.1 with the additional restrictions that there exists $M > 0$ such that σ is bounded by M and that there exists $\beta \in (0,1]$ such that for any $s, s' \in [0,1]$, for any $y \in \mathbb{R}^d$ and for all $\mu \in \mathcal{P}_2\left(\mathbb{R}^d\right)$, we have:

$$|\sigma(s,y,\mu) - \sigma(s',y,\mu)| \leq L|s-s'|^\beta, \quad |b(s,y,\mu) - b(s',y,\mu)| \leq L|s-s'|^\beta.$$

We additionally assume that b_ε and σ_ε have adequate conditions to ensure the existence and uniqueness of a solution to the McKean-Vlasov Equation.

Theorem 7.2.2. *Let $\alpha \in (0,1/2)$. Let A be a Borel set of the space of $\mathbb{R}^d$-valued continuous paths over $[0,T]$ in the Hölder topology of $C^\alpha([0,T];\mathbb{R}^d)$. Let*

$$\mathcal{I}(A) := \inf_{h\in\mathcal{H}} \Big\{ \tfrac{\|h\|^2_{\mathcal{H}}}{2}; \Phi(h) \in A \Big\}.$$

Then

$$-\mathcal{I}(\mathring{A}) \leq \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\Big[X^\varepsilon \in A\Big] \leq \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\Big[X^\varepsilon \in A\Big] \leq -\mathcal{I}(\bar{A}),$$

where $\mathring{A}$ and $\bar{A}$ are the interior and closure of the set A with respect to the topology generated by the Hölder norm.

In order to prove the Theorem 7.2.2 we first prove the following asymptotic result:

Proposition 7.2.3. *Let $h \in \mathcal{H}$. Take $\forall R, \rho > 0$, $\exists \delta, \nu > 0$ such that $\forall 0 < \varepsilon < \nu$,*

$$\mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta\Big] \lesssim \exp\Big(-\frac{R}{\varepsilon}\Big).$$

Intuitively, Proposition 7.2.3 quantifies the probability of a highly varying process in $\|\cdot\|_\alpha$ when the equation's input signal is small in $\|\cdot\|_\infty$.

7.3 Proof of Proposition 7.2.3

Proof of Proposition 7.2.3. Let $t \in [0,T]$. Fix $R, \rho > 0$. In order to progress with a Local Lipschitz condition, we first need to consider the function $\Phi(h)$ (recall (7.2.1)) for $h \in \mathcal{H}$. This is a continuous solution of an ODE on the compact interval $[0,T]$. Therefore, it is bounded and we can say that $\exists N > 0$ such that $\|\Phi(h)\|_\infty < N$.

We condition on the event that the process X^ε remains in the ball of radius N and we see

$$\begin{aligned}&\mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta\Big]\\ &\leq \mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta, \|X^\varepsilon\|_\infty < N\Big] + \mathbb{P}\Big[\|X^\varepsilon\|_\infty \geq N\Big].\end{aligned}$$

We use that we have the LDP result for X^ε in a supremum norm and choose N large enough so that

$$\mathbb{P}\Big[\|X^\varepsilon\|_\infty \geq N\Big] < \exp\Big(-\frac{R}{\varepsilon}\Big).$$

We also introduce a step function approximation to discretize the process X^ε in (7.0.1) as, for $l \in \mathbb{N}$

$$X^{\varepsilon,l}_t = X^\varepsilon_{\frac{j}{l}} \text{ on the interval } t \in \Big(\tfrac{j}{l}, \tfrac{j+1}{l}\Big], \quad \text{with } X^{\varepsilon,l}_0 = x.$$

Step 1. Analysis of the diffusion term for $h = 0$. Consider

$$\mathbb{P}\Big[\|\sqrt{\varepsilon}\int_0^{\cdot}\sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})dW_s\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W\|_\infty \leq \delta, \|X^\varepsilon\|_\infty < N\Big] \tag{7.3.1}$$

$$\leq \mathbb{P}\Big[\|\sqrt{\varepsilon}\int_0^{\cdot}[\sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon}) - \sigma_\varepsilon(\tfrac{\lfloor sl\rfloor}{l}, X^{\varepsilon,l}, \mathcal{L}^{X^\varepsilon}_{\frac{\lfloor sl\rfloor}{l}})]dW_s\|_\alpha \geq \frac{\rho}{2},$$

$$\tfrac{1}{l^\beta} + \|X^\varepsilon - X^{\varepsilon,l}\|_\infty + \mathbb{E}[\|X^\varepsilon - X^{\varepsilon,l}\|_\infty^2]^{1/2} \leq \gamma\Big] \tag{7.3.2}$$

$$+\,\mathbb{P}\Big[\tfrac{1}{l^\beta} + \|X^\varepsilon - X^{\varepsilon,l}\|_\infty + \mathbb{E}[\|X^\varepsilon - X^{\varepsilon,l}\|_\infty^2]^{1/2} > \gamma, \|X_\varepsilon^x\|_\infty < N\Big] \tag{7.3.3}$$

$$+\,\mathbb{P}\Big[\|\sqrt{\varepsilon}\int_0^{\cdot}\sigma_\varepsilon(\tfrac{\lfloor sl\rfloor}{l}, X_s^{\varepsilon,l}, \mathcal{L}^{X^\varepsilon}_{\frac{\lfloor sl\rfloor}{l}})dW_s\|_\alpha \geq \frac{\rho}{2}, \|\sqrt{\varepsilon}W\|_\infty \leq \delta\Big]. \tag{7.3.4}$$

Firstly, consider the term (7.3.2). We denote

$$\eta_\varepsilon = \sup_{s,y,\mu}\Big\{\,|b(s,y,\mu) - b_\varepsilon(s,y,\mu)|\,,\,|\sigma(s,y,\mu) - \sigma_\varepsilon(s,y,\mu)|\Big\}.$$

By uniform convergence of b_ε to b and σ_ε to σ, we have that $\lim_{\varepsilon\to 0}\eta_\varepsilon = 0$. We choose ε small enough so that $\eta_\varepsilon \leq \frac{L\gamma}{4}$. Then

$$\begin{aligned}(7.3.2) \leq &\mathbb{P}\Big[\Big\|\int_0^{\cdot}\frac{2[\sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon}) - \sigma_\varepsilon(\frac{\lfloor sl\rfloor}{l}, X_s^{\varepsilon,l}, \mathcal{L}^{X^\varepsilon}_{\frac{\lfloor sl\rfloor}{l}})]}{L\gamma}dW_s\Big\|_\alpha \geq \frac{\rho}{\sqrt{\varepsilon}L\gamma},\\ &\qquad\frac{2\|\sigma_\varepsilon(\cdot, X^\varepsilon, \mathcal{L}^{X^\varepsilon}) - \sigma_\varepsilon(\frac{\lfloor \cdot l\rfloor}{l}, X^{\varepsilon,l}, \mathcal{L}^{X^{\varepsilon,l}})\|_\infty}{L\gamma} \leq 1\Big]\\ &\leq C'\exp\Big(\frac{-\rho^2}{C'L^2\gamma^2\varepsilon}\Big),\end{aligned}$$

using Lemma 7.1.1. Thus choose γ such that $\frac{\rho^2}{C'L^2R} \geq \gamma^2$.

Secondly, consider the term (7.3.3). We take ε small enough so that $\eta_\varepsilon < 1$. Applying Itô's formula to $|X_\varepsilon^x(t)|^2$ gives

$$\begin{aligned}|X_t^\varepsilon|^2 = |x|^2 + 2\sqrt{\varepsilon}\int_0^t\Big\langle X_s^\varepsilon, \sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})dW_s\Big\rangle + 2\int_0^t\Big\langle X_s^\varepsilon, b_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})\Big\rangle ds\\ + \varepsilon\int_0^t \mathrm{Tr}\Big(\sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})\sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})^T\Big)ds.\end{aligned}$$

Following the estimation methods used to prove Theorem 6.1.2 we have

$$\mathbb{E}\Big[\|X^\varepsilon\|_\infty^2\Big] \leq K\Big(|x|^2 + \mathbb{E}\Big[\|b(\cdot,0,\delta_0)\|_\infty^2\Big]\Big)\exp\Big(K + \mathbb{E}\Big[\|b(\cdot,0,\delta_0)\|_\infty^2\Big]\Big) < \infty.$$

In the same way, we can additionally prove $\mathbb{E}[\,\|X_\varepsilon^x\|_\infty^{2q}] < \infty$. Let $j = \lfloor tl\rfloor$. We can rewrite our SDE, for $t \in [\frac{j}{l}, \frac{j+1}{l}]$, as

$$X^\varepsilon_{\frac{j}{l},t} = \sqrt{\varepsilon}\int_{\frac{j}{l}}^t\sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})dW_s + \int_{\frac{j}{l}}^t b_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})ds.$$

Taking expectations yields

$$
\begin{aligned}
&\mathbb{E}\Big[\|X^\varepsilon - X^{\varepsilon,l}\|_\infty^2\Big] \\
&\le 2\mathbb{E}\Big[\sup_j \sup_{t\in\left[\frac{j}{l},\frac{j+1}{l}\right]} \Big|\int_{\frac{j}{l}}^t b_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})ds\Big|^2\Big] + 2\mathbb{E}\Big[\sup_j \sup_{t\in\left[\frac{j}{l},\frac{j+1}{l}\right]} \Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_s\Big|^2\Big] \\
&\le \frac{1}{l^2}\Big(4\eta_\varepsilon + 32L^2\mathbb{E}\Big[\|X^\varepsilon\|_\infty^{2q}\Big] + 8L^2 4\mathbb{E}\Big[\|X^\varepsilon\|_\infty^2\Big] + 4\|b(\cdot,0,\delta_0)\|_\infty^2\Big) + \frac{8\varepsilon M^2}{l} \lesssim \frac{1}{l},
\end{aligned}
$$

and we write $\mathbb{E}\big[\|X^\varepsilon - X^{\varepsilon,l}\|_\infty^2\big]^{1/2} \le \frac{K_1}{\sqrt{l}}$. In the same way we also have that

$$
\begin{aligned}
\|X^\varepsilon - X^{\varepsilon,l}\|_\infty &= \sup_{j=0,\dots,l-1}\sup_{t\in[\frac{j}{l},\frac{j+1}{l}]}\Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_r\Big| \\
&\qquad + \sup_{j=0,\dots,l-1}\sup_{t\in[\frac{j}{l},\frac{j+1}{l}]}\Big|\int_{\frac{j}{l}}^t b_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dr\Big| \\
&\le \sup_{j,t}\Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_s\Big| \\
&\qquad + \sup_{j,t}\int_{\frac{j}{l}}^t \Big[\eta_\varepsilon + L(1+\|X\|_\infty^q) + L\mathbb{E}\big[\|X\|_\infty^2\big]^{1/2} + \sup_{r\in[0,T]}\Big|b(r,0,\delta_0)\Big|\Big]ds \\
&\le \sup_{j,t}\Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_s\Big| + K_2\frac{1+\|X^\varepsilon\|_\infty^q}{l}.
\end{aligned}
$$

Hence we have for term (7.3.3) that

$$
\begin{aligned}
&\mathbb{P}\Big[\tfrac{1}{l^\beta} + \|X^\varepsilon - X^{\varepsilon,l}\|_\infty + \mathbb{E}[\|X^\varepsilon - X^{\varepsilon,l}\|_\infty^2]^{1/2} > \gamma, \|X^\varepsilon\|_\infty < N\Big] \\
&\quad\le \mathbb{P}\Big[\sup_{j,t}\Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_s\Big| \\
&\qquad\qquad + K_2\frac{1+\|X^\varepsilon\|_\infty^q}{l} + \frac{K_1}{\sqrt{l}} + \frac{1}{l^\beta} > \gamma, \|X^\varepsilon\|_\infty < N\Big] \\
&\quad\le \mathbb{P}\Big[\sup_{j,t}\Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_s\Big| > \gamma - \frac{K_3}{l^{\frac{1}{2}\wedge\beta}}\Big],
\end{aligned}
$$

where $K_3 = K_1 + K_2(1+N^q) + 1$.

Therefore, using Chernoff's inequality

$$
\begin{aligned}
(7.3.3) &\le \mathbb{P}\Big[\sup_{j,t}\exp\Big(\lambda\Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_s\Big|\Big) > \exp\Big(\frac{\lambda}{l^{\frac{1}{2}\wedge\beta}}\big(\gamma l^{\frac{1}{2}\wedge\beta} - K_3\big)\Big)\Big] \\
&\le \frac{\sup_{j,t}\mathbb{E}\Big[\exp\Big(\lambda\Big|\sqrt{\varepsilon}\int_{\frac{j}{l}}^t \sigma_\varepsilon(s,X_s^\varepsilon,\mathcal{L}_s^{X^\varepsilon})dW_s\Big|\Big)\mathbb{1}_{\|X^\varepsilon\|_\infty<N}\Big]}{\exp\Big(\frac{\lambda}{l^{\frac{1}{2}\wedge\beta}}(\gamma l^{\frac{1}{2}\wedge\beta} - K_3)\Big)} \\
&\lesssim \exp\Big(\lambda^2\varepsilon\frac{M}{2l} - \frac{\lambda}{l^{\frac{1}{2}\wedge\beta}}(\gamma l^{\frac{1}{2}\wedge\beta} - K_3)\Big) \lesssim \exp\Big(\frac{-(\gamma l^{\frac{1}{2}\wedge\beta} - K_3)^2}{2\varepsilon M}\frac{l}{l^{\frac{1}{2}\wedge\beta}}\Big),
\end{aligned}
$$

by optimizing over the arbitrary choice of λ. We can now choose the constant l such that $\frac{(\gamma l^{\frac{1}{2}\wedge\beta} - K_3)^2}{2M} l^{1-(1\wedge 2\beta)} > R$.

Finally, to evaluate Equation (7.3.4), we first consider $\sigma_\varepsilon(\cdot, X_\cdot^{\varepsilon,l}, \mathcal{L}^{X^\varepsilon}_{\frac{\lfloor\cdot l\rfloor}{l}})$. This process is constant over the interval $(\frac{j}{l}, \frac{j+1}{l}]$. Then taking the Hölder norm we get

$$\begin{aligned}
&\Big\| \int_0^\cdot \sigma_\varepsilon(\tfrac{\lfloor ls \rfloor}{l}, X_s^{\varepsilon,l}, \mathcal{L}^{X^\varepsilon}_{\frac{\lfloor sl \rfloor}{l}}) dW_s \Big\|_\alpha \\
&\quad = \Big\| \sum_{j=0}^{l-1} \sigma_\varepsilon\big(\tfrac{j}{l}, X^{\varepsilon,l}_{\frac{j}{l}}, \mathcal{L}^{X^\varepsilon}_{\frac{j}{l}}\big)\big[W_{\frac{j+1}{l}\wedge \cdot} - W_{\frac{j}{l}\wedge\cdot}\big] \Big\|_\alpha \le 2lM\|W\|_\alpha,
\end{aligned}$$

using $\|\sigma_\varepsilon\|_\infty \le M$ and the triangle inequality. Then

$$\begin{aligned}
&\mathbb{P}\Big[\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X_s^{\varepsilon,l}, \mathcal{L}_s^{X^{\varepsilon,l}}) dW_s\|_\alpha \ge \frac{\rho}{2}, \|\sqrt{\varepsilon}W\|_\infty \le \delta\Big] \\
&\qquad \le \mathbb{P}\Big[\|W\|_\alpha \ge \frac{\rho}{4\sqrt{\varepsilon}lM}, \|W\|_\infty \le \frac{\delta}{\sqrt{\varepsilon}}\Big] \\
&\qquad \le C\max\Big(1, \Big(\frac{\rho}{4Ml\delta}\Big)^{1/\alpha}\Big)\exp\Big(\frac{-1}{\varepsilon}\frac{1}{C}\Big(\frac{\rho}{4Ml\delta}\Big)^{1/\alpha}\delta^2\Big),
\end{aligned}$$

where we applied Lemma 7.1.2 and chose δ such that $\frac{\rho}{R^\alpha 4MlC^\alpha} \ge \delta^{1-2\alpha}$.

Injecting these three results in (7.3.1) gives us that

$$\mathbb{P}\Big[\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}^{X^\varepsilon}_{\frac{\lfloor sl \rfloor}{l}}) dW_s\|_\alpha \ge \rho, \|\sqrt{\varepsilon}W\|_\infty \le \delta, \|X^\varepsilon\|_\infty < N\Big] \lesssim \exp\Big(-\frac{R}{\varepsilon}\Big). \tag{7.3.5}$$

Step 2. The Hölder norm of the whole process when $h = 0$. We have

$$\begin{aligned}
\|X^\varepsilon - \Phi(0)\|_{\alpha,t} \le& \big\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon}) dW_s\big\|_{\alpha,t} &(7.3.6)\\
&+ \big\|\int_0^\cdot [b_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon}) - b(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})] ds\big\|_{\alpha,t} &(7.3.7)\\
&+ \big\|\int_0^\cdot [b(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon}) - b(s, \Phi(0)_s, \delta_{\Phi(0)_s})] ds\big\|_{\alpha,t}. &(7.3.8)
\end{aligned}$$

Equation (7.3.6) is the term in (7.3.5) that we desire. Equation (7.3.7) is bounded above by η_ε. We only consider the cases when $\|X^\varepsilon\|_\infty, \|\Phi(0)\|_\infty < N$ since we know that $\Phi(0)_t$ remains in this ball and we conditioned on X_t^ε remaining in the same ball. This means that by the Locally Lipschitz condition, we can say that $b(t, x, \mu)$ is Lipschitz in the spacial variable with constant L_N. Therefore for (7.3.8) we have

$$\begin{aligned}
&\Big\|\int_0^\cdot b(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon}) - b(s, \Phi(0)_s, \delta_{\Phi(0)_s}) ds\Big\|_{\alpha,t} \\
&\quad \le \sup_{p,q\in[0,t]} \frac{\int_p^q |b(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon}) - b(s, \Phi(0)_s, \delta_{\Phi(0)_s})|}{|q-p|^\alpha} \\
&\quad \le \sup_{p,q\in[0,t]} \frac{L_N}{|q-p|^\alpha}\int_p^q |X_s^\varepsilon - \Phi(0)_s| ds + \frac{L}{|q-p|^\alpha}\int_p^q \mathbb{E}\Big[|X_s^\varepsilon - \Phi(0)_s|^2\big]^{1/2} ds \\
&\quad \le L_N\|X^\varepsilon - \Phi(0)\|_{\infty,t} + L_N\int_0^t \|X^\varepsilon - \Phi(0)\|_{\alpha,s} ds &(7.3.9)\\
&\qquad + L\mathbb{E}\big[\|X^\varepsilon - \Phi(0)\|_\infty^2\big]^{1/2}, &(7.3.10)
\end{aligned}$$

using Lemma 6.0.1.

Next, we want to show that the strong error $\mathbb{E}\big[\|X_\varepsilon^x - \Phi(0)\|_\infty^2\big]$ can be controlled by ε. Using that

$$\begin{aligned} d\big(X^\varepsilon - \Phi(0)\big)_t =&\sigma_\varepsilon\Big(t, X^\varepsilon_t, \mathcal{L}^{X^\varepsilon}_t\Big)dW_t \\ &+ \Big(b_\varepsilon(t, X^\varepsilon_t, \mathcal{L}^{X^\varepsilon}_t) - b(t, X^\varepsilon_t, \mathcal{L}^{X^\varepsilon}_t)\Big)dt \\ &+ \Big(b(t, X^\varepsilon_t, \mathcal{L}^{X^\varepsilon}_t) - b(t, \Phi(0)_t, \delta_{\Phi(0)_t})\Big)dt, \end{aligned}$$

and Itô's formula for $f(x) = |x|^2$ with $X^\varepsilon_0 - \Phi(0)_0 = 0$ gives

$$\begin{aligned} \|X^\varepsilon - \Phi(0)\|^2_{\infty,t} \le& 2\sqrt{\varepsilon} \sup_{0\le s\le t} \Big| \int_0^s \langle X^\varepsilon_r - \Phi(0)_r, \sigma_\varepsilon(r, X^\varepsilon_r, \mathcal{L}^{X^\varepsilon}_r)dW_r\rangle \Big| \\ &+ \varepsilon M^2 td + \int_0^t 2\eta_\varepsilon |X^\varepsilon_r - \Phi(0)_r| dr \\ &+ 2\int_0^t \Big| \langle X^\varepsilon_r - \Phi(0)_r, b(r, X^\varepsilon_r, \mathcal{L}^{X^\varepsilon}_r) - b(r, \Phi(0)_r, \delta_{\Phi(0)_r})\rangle \Big| dr. \end{aligned}$$

Taking expectations gives

$$\begin{aligned} \mathbb{E}\Big[\|X^\varepsilon - \Phi(0)\|^2_{\infty,t}\Big] \le& \mathbb{E}\Big[\sup_{s\in[0,t]} 2\sqrt{\varepsilon}\int_0^s \langle X^\varepsilon_r - \Phi(0)_r, \sigma_\varepsilon(r, X^\varepsilon_r, \mathcal{L}^{X^\varepsilon}_r)dW_r\rangle\Big] + \varepsilon M^2 t \\ &+ 2L\int_0^t \mathbb{E}\Big[\|X^\varepsilon - \Phi(0)\|^2_{\infty,s}\Big]ds + \int_0^t 2\eta_\varepsilon \mathbb{E}\Big[\|X^\varepsilon - \Phi(0)\|_{\infty,s}\Big]ds \\ &+ 2L\int_0^t \mathbb{E}\Big[\|X^\varepsilon - \Phi(0)\|^2_{\infty,s}\Big]^{1/2}\mathbb{E}\Big[\|X^\varepsilon - \Phi(0)\|_{\infty,s}\Big]ds \\ \le& \varepsilon(C_1 + Mt) + (C_1M^2 + 4L + 1)\int_0^t \mathbb{E}\Big[\|X^\varepsilon - \Phi(0)\|^2_{\infty,s}\Big]ds + \eta^2_\varepsilon \end{aligned}$$

Refining, we then obtain $\mathbb{E}\big[\|X^x_\varepsilon - \Phi^x(0)\|^2_\infty\big]^{1/2} \le K(\eta_\varepsilon \vee \varepsilon^{1/2})e^K$. We have shown that this expectation is of order $\varepsilon^{1/2}$. Now we consider $\|X^x_\varepsilon - \Phi^x(0)\|_\infty$. Since the supremum norm can be made to appear inside the integrals, we have

$$\begin{aligned} \|X^\varepsilon - \Phi(0)\|_{\infty,t} \le& \|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X^\varepsilon_s, \mathcal{L}^{X^\varepsilon}_s)dW_s\|_{\infty,t} + \eta_\varepsilon t \\ &+ \int_0^t L_N\|X^\varepsilon - \Phi(0)\|_{\infty,r}dr + Lt\mathbb{E}[\|X^\varepsilon - \Phi(0)\|^2_\infty]^{1/2}, \end{aligned}$$

and by using Grönwall, we get

$$\begin{aligned} \|X^\varepsilon - \Phi(0)\|_{\infty,t} \le& \bigg(\Big\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X^\varepsilon_s, \mathcal{L}^{X^\varepsilon}_s)dW_s\Big\|_{\infty,t} + \eta_\varepsilon + \mathbb{E}\Big[\|X^\varepsilon - \Phi(0)\|^2_\infty\Big]^{1/2}\bigg)e^{L_N t} \\ \le& \bigg(\Big\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X^\varepsilon_s, \mathcal{L}^{X^\varepsilon}_s)dW_s\Big\|_{\alpha,t} + (\sqrt{\varepsilon}\vee\eta_\varepsilon)K'\bigg)e^{K't}. \end{aligned} \tag{7.3.11}$$

Combining Equation (7.3.9), Equation (7.3.10) and Equation (7.3.11) gives

$$\begin{aligned} \Big\|X^\varepsilon - \Phi(0)\Big\|_\alpha \le& \Big\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X^\varepsilon_s, \mathcal{L}^{X^\varepsilon}_s)dW_s\Big\|_\alpha\Big(1 + L_N e^{K'}\Big)e^{L_N} \\ &+ (\eta_\varepsilon \vee \sqrt{\varepsilon})\Big(1 + L_N K' e^{K'} + KLe^K\Big)e^{L_N} \\ \le& \bigg[\Big\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X^\varepsilon_s, \mathcal{L}^{X^\varepsilon}_s)dW_s\Big\|_\alpha + (\eta_\varepsilon \vee \sqrt{\varepsilon})\bigg]K_4. \end{aligned}$$

Thus for any choice of ρ we see that

$$\mathbb{P}\Big[\|X^\varepsilon - \Phi(0)\|_\alpha \geq \rho, \|X^\varepsilon\|_\infty < N\Big]$$
$$\leq \mathbb{P}\Big[\Big(\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})dW_s\|_\alpha + (\eta_\varepsilon \vee \sqrt{\varepsilon})\Big)K_4 \geq \rho, \|X^\varepsilon\|_\infty < N\Big].$$

and by choosing ε small enough such that $\eta_\varepsilon \vee \sqrt{\varepsilon} < \frac{\rho}{2K_4}$ we get

$$\begin{aligned}
&\mathbb{P}\Big[\|X^\varepsilon - \Phi(0)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W\|_\infty \leq \delta, \|X^\varepsilon\|_\infty < N\Big]\\
&\leq \mathbb{P}\left[\Big\|\sqrt{\varepsilon}\int_0^\cdot \sigma_\varepsilon(s, X_s^\varepsilon, \mathcal{L}_s^{X^\varepsilon})dW_s\Big\|_\alpha \geq \frac{\rho}{2K_4}, \Big\|\sqrt{\varepsilon}W\Big\|_\infty \leq \delta, \|X^\varepsilon\|_\infty < N\right]\\
&\lesssim \exp\Big(-\frac{R}{\varepsilon}\Big),
\end{aligned}$$

since in Equation (7.3.5) the choice of ρ is arbitrary.

Step 3. The case when $h \neq 0$. For the final step, we use the same method as in [BAL94] to extend the results to Wiener processes with drift. Using a Girsanov transformation we have that there is a measure $\tilde{\mathbb{P}}$ absolutely continuous to the standard probability measure $\mathbb{P}$.

Note that the law of the stochastic process is not changed by perturbing the path of the Brownian motion by some element of the Cameron Martin space. When solving a McKean-Vlasov equation (unlike classical SDEs), one has to fix the law of the probability space in order to define $\mathcal{L}^X = \mathbb{P} \circ X^{-1}$. Hence the law is not changed when one considers a different driving noise for the SDE. This is most obvious in expression (7.2.1) where the delta distribution follows the path of the skeleton with input $h = 0$.

We rewrite the SDE and skeleton process

$$\begin{aligned}
X_t^\varepsilon =& x + \int_0^t b_\varepsilon\Big(s, X_s^\varepsilon, \mathbb{P}\circ[X_s^\varepsilon]^{-1}\Big)ds + \int_0^t \sigma_\varepsilon\Big(s, X_s^\varepsilon, \mathbb{P}\circ[X_s^\varepsilon]^{-1}\Big)\dot{h}_s ds\\
&+ \sqrt{\varepsilon}\int_0^t \sigma_\varepsilon\Big(s, X_s^\varepsilon, \mathbb{P}\circ[X_s^\varepsilon]^{-1}\Big)d\tilde{W}_s,\\
\Phi(h)_t =& x + \int_0^t b\Big(s, \Phi(h)_s, \delta_{\Phi(0)_s}\Big)ds + \int_0^t \sigma\Big(s, \Phi(h)_s, \delta_{\Phi(0)_s}\Big)\dot{h}_s ds,
\end{aligned}$$

where $\tilde{W} = W - h/\sqrt{\varepsilon}$, $\tilde{\mathbb{P}}$ is the measure where $\tilde{W}$ is a Brownian motion and $\mathbb{P}\circ[X_t^\varepsilon]^{-1} = \mathcal{L}_t^{X^\varepsilon}$. The drift term $b_\varepsilon + \sigma_\varepsilon \dot{h}$ satisfies the properties from before and matches the skeleton process $\Phi(h)$.

Also note that

$$\mathbb{W}^{(2)}\Big(\mathbb{P}\circ[X_t^\varepsilon]^{-1}, \delta_{\Phi(0)_t}\Big) = \mathbb{E}^{\mathbb{P}}\Big[|X_t^\varepsilon - \Phi(0)_t|^2\Big]^{1/2},$$

which we have already showed to go to 0 as $\varepsilon \to 0$. Thus we argue in the same way as in Step 2 and conclude

$$\begin{aligned}
&\mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta\Big]\\
&\quad\lesssim \tilde{\mathbb{P}}\Big[\|\tilde{X^\varepsilon} - \Phi(0)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}\tilde{W}\|_\infty \leq \delta\Big] \lesssim \exp\Big(-\frac{R}{\varepsilon}\Big).
\end{aligned}$$

□

7.4 Proof of Theorem 7.2.2

We are now in position to prove our main result, Theorem 7.2.2.

Proof of Theorem 7.2.2. Proving the upper bound. First consider the case where $0 \notin A$ and A is closed in the Hölder Topology. Then there exists an r such that $\mathcal{I}(A) > r > 0$. Let us consider the ball in the Cameron-Martin space $\mathcal{H}$

$$\Big\{h \in \mathcal{H} : \tfrac{\|h\|_{\mathcal{H}}^2}{2} \leq r\Big\}.$$

Recall that if $h \in \mathcal{H}$ then $h \in C^{1/2}([0,T];\mathbb{R}^{d'})$ and is bounded and, moreover, that $\|h\|_\infty \leq \|h\|_{\frac{1}{2}-\text{Höl}} \leq \|h\|_{\mathcal{H}}$. Therefore we can apply Arzelà-Ascoli Theorem [DS88] to get that this set is compact. Hence we can find a finite open cover of this set and we can restrict the radius of the open balls. We write

$$\Big\{h \in \mathcal{H} : \tfrac{\|h\|_{\mathcal{H}}^2}{2} \leq r\Big\} \subset \bigcup_{i=1}^{N} B_\infty(h_i, \eta_{h_i}) = U.$$

These balls are in the uniform topology and the elements h_i are all have $\|h\|_{\mathcal{H}} < \sqrt{2r}$. By this property, $\Phi(h_i) \notin A$. If it were, $\|h\|_{\mathcal{H}}^2 > 2\mathcal{I}(A)$. The set A is closed in the Hölder topology so A^c is open in the Hölder topology. Therefore, there exists a ρ_{h_i} such that in the Hölder topology $B_\alpha(\Phi(h_i), \rho_{h_i})$ is in A^c, and therefore does not intersect with A. Hence when $X^\varepsilon \in A$, we can say that $\|X^\varepsilon - \Phi(h_i)\|_\alpha \geq \rho_{h_i}$. Finally, we can estimate

$$\begin{aligned}
\mathbb{P}[X^\varepsilon \in A] &= \mathbb{P}[X^\varepsilon \in A, \sqrt{\varepsilon}W \notin U] + \mathbb{P}[X^\varepsilon \in A, \sqrt{\varepsilon}W \in U] \\
&\leq \mathbb{P}[\sqrt{\varepsilon}W \notin U] + \mathbb{P}[X^\varepsilon \in A, \sqrt{\varepsilon}W \in U] \\
&\leq \mathbb{P}[\sqrt{\varepsilon}W \notin U] + \sum_{i=1}^{N} \mathbb{P}\Big[\|X^\varepsilon - \Phi(h_i)\|_\alpha \geq \rho_{h_i}, \|\sqrt{\varepsilon}W - h_i\|_\infty \leq \eta_{h_i}\Big] \\
&\leq \mathbb{P}\Big[\sqrt{\varepsilon}W \notin U\Big] + N \exp\Big(-\frac{2r}{\varepsilon}\Big),
\end{aligned}$$

where for the last line we apply Proposition 7.2.3, with η_{h_i} and ε chosen sufficiently small for the given ρ_{h_i}. The η_{h_i} are dependent on our choice of open cover for the compact set, so we can make them as small as required. We already have a LDP for a Wiener process on the uniform norm by [HIPP14]. Hence we have for ε sufficiently small that

$$\mathbb{P}\Big[\sqrt{\varepsilon}W \notin U\Big] \leq \exp\Big(-\tfrac{\mathcal{I}(U^c)}{\varepsilon}\Big).$$

If $h \notin U$, then we have that $\|h\|_{\mathcal{H}}^2 > 2r$ and consequently $\mathbb{P}\Big[\sqrt{\varepsilon}W \notin U\Big] \leq \exp\Big(-2r/\varepsilon\Big)$. Combining all of this together we get

$$\limsup_{\varepsilon \to 0} \varepsilon \log\Big(\mathbb{P}\Big[X^\varepsilon \in A\Big]\Big) \leq -r,$$

where r was chosen arbitrarily so that $r < \mathcal{I}(A)$ where A is closed. We optimize for our choice of r and get

$$\limsup_{\varepsilon \to 0} \varepsilon \log\Big(\mathbb{P}\Big[X^\varepsilon \in A\Big]\Big) \leq -\mathcal{I}(A),$$

which is the upper inequality for the Theorem.

Proving the lower bound. Now consider A to be an open set in the Hölder topology and let $h \in \mathcal{H}$ such that $\Phi(h) \in A$. There exists a $\rho > 0$ such that the Hölder ball $B_\alpha(\Phi(h), \rho) \subset A$. Also we have that

$$\mathbb{P}\Big[\|\sqrt{\varepsilon}W - h\|_\infty < \delta\Big] \leq \mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta\Big] + \mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha < \rho\Big].$$

Hence

$$\begin{aligned}
\mathbb{P}\Big[X^\varepsilon \in A] \geq& \mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha < \rho\Big] \\
\geq& \mathbb{P}\Big[\|\sqrt{\varepsilon}W - h\|_\infty < \delta\Big] - \mathbb{P}\Big[\|X^\varepsilon - \Phi(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta\Big] \\
\geq& \mathbb{P}\Big[\|\sqrt{\varepsilon}W - h\|_\infty < \delta\Big] - \exp\Big(-\frac{R}{\varepsilon}\Big).
\end{aligned}$$

Applying the LDP for the Brownian motion (see [DZ10, Theorem 5.2.23]) and using that $\frac{\|h\|^2_{\mathcal{H}}}{2} \geq \mathcal{I}(B_\infty(h,\delta))$, we see that

$$\mathbb{P}\Big[\|\sqrt{\varepsilon}W - h\|_\infty < \delta\Big] \geq \exp\Big(-\frac{\|h\|^2_{\mathcal{H}}}{2\varepsilon}\Big)$$

and hence

$$\mathbb{P}\Big[X^\varepsilon \in A\Big] \geq \exp\Big(-\tfrac{\|h\|^2_{\mathcal{H}}}{2\varepsilon}\Big) - \exp\Big(-\tfrac{R}{\varepsilon}\Big),$$

where we can choose R to take any value. Choosing $R = \|h\|^2_{\mathcal{H}}$ and rearranging we get

$$\mathbb{P}\Big[X^\varepsilon \in A\Big] \geq \exp\Big(-\tfrac{\|h\|^2_{\mathcal{H}}}{2\varepsilon}\Big)\Big(1 - \exp\Big(-\tfrac{\|h\|^2_{\mathcal{H}}}{2\varepsilon}\Big)\Big).$$

Hence

$$\liminf_{\varepsilon\to 0} \varepsilon \log\Big(\mathbb{P}\Big[X^\varepsilon \in A\Big]\Big) \geq \liminf_{\varepsilon\to 0} \varepsilon \log\Big(1 - \exp\Big(-\tfrac{\|h\|^2_{\mathcal{H}}}{2\varepsilon}\Big)\Big) - \tfrac{\|h\|^2_{\mathcal{H}}}{2}.$$

The limit goes to 0 for any choice of $h \in \mathcal{H}$. Finally, as h was arbitrarily chosen in A, we take the infimum over h and get

$$\liminf_{\varepsilon\to 0} \varepsilon \log\Big(\mathbb{P}\Big[X^\varepsilon \in A\Big]\Big) \geq -\mathcal{I}(A).$$

This completes the proof of the Theorem. □

Chapter 8

Functional Iterated Logarithm Law

In this Chapter, we apply Theorem 7.2.2 with a specific choice of rescaling with the goal of observing the Functional Iterated Logarithm law. The results of this Chapter can be found published in [dRST19, Section 5].

8.1 Strassen's Law for Brownian motion

Strassen's Law, or the Law of Iterated Logarithm describes the magnitude of the fluctuations of a Brownian motion. It was first proved in [Str64]. Observe that for a Brownian Motion W_t, we have that $X_n^{(1)}(t) = W_{nt}/n \to 0$ both in probability and almost surely as $n \to \infty$. However $X_n^{(2)}(t) = W_{nt}/\sqrt{n}$ is a Brownian Motion for any choice of n. Therefore, something is happening between n and $\sqrt{n}$ which is turning a stochastic process into a deterministic constant in the limit as $n \to \infty$. Strassen's Law says that

$$X_n^{(3)}(t) = \frac{W_{nt}}{\sqrt{n \log\log(n)}},$$

converges to 0 in probability but does not converge almost surely. In particular

$$\limsup_{n\to\infty} X_n^{(3)}(1) = \sqrt{2}, \quad \text{almost surely.}$$

In this section we are interested in studying whether stochastic processes have a similar type of property. We will consider the solution of the SDE run over a large time interval of order n and rescaled to order $\sqrt{n \log(\log(n))}$. Similar to the proof of Strassen's Law, we will show that the set of rescaled paths is relatively compact in the Hölder topology but that the set of limit points of this set is uncountable which implies the failure of almost sure convergence.

In [Bal86], Baldi proves a Law of Iterated Logarithm for classical SDEs for the uniform topology. This was then extended in [Edd00] and later [EN02] to other coarser pathspace topologies. Standard LDP results easily give us convergence in probability. We calculate the set of possible limit points of the scaled diffusions which for a classical SDE are

$$\Big\{ \Phi(h) : d\Phi(h)_t = b\big(\Phi(h)_t\big)dt + \sigma\big(\Phi(h)_t\big)dh_t, \quad \Phi(h)_0 = x, \|h\|_{\mathcal{H}} \leq \sqrt{2} \Big\}.$$

We show below, that similarly for a McKean-Vlasov SDE these are

$$\Big\{ \Phi(h) : d\Phi(h)_t = b\big(\Phi(h)_t, \delta_{\Phi(0)_t}\big)dt + \sigma\big(\Phi(h)_t, \delta_{\Phi(0)_t}\big)dh_t, \quad \Phi(h)_0 = x, \|h\|_{\mathcal{H}} \leq \sqrt{2} \Big\}.$$

We will follow the methods of [Bal86], [Edd00] and [EN02] to extend the LDP results to prove an Iterated Logarithm Law for the class of McKean-Vlasov SDEs in Theorem 6.1.2. It seems possible to use microscopic rescaling of the Brownian motion such as in [Gan93] to provide an alternative proof of our result, however, we do not pursue this point.

Remark 8.1.1 (Decoupling Argument). *To the best of our knowledge, there are no results proving a Strassen type law for SDEs with coefficients which can vary in time and we were unable to establish any such results while working on this paper. The conditions that we require on the measure dependency are similar to those of the spacial dependency and do not naturally translate into conditions for time dependency. Therefore, proving that they are satisfied is much easier in the McKean-Vlasov Equation setting when they are written as properties on the measure dependency than for some general time dependent coefficient.*

8.2 Functional Iterated Logarithm Law for McKean-Vlasov Equations

Firstly, we need to define in what sense we will be rescaling our McKean-Vlasov Equations.

Definition 8.2.1. *Let $\alpha \in \mathbb{R}^+$. A family of continuous bijections $\Gamma_\alpha : \mathbb{R}^d \to \mathbb{R}^d$ is said to be a System of Contractions centered at x if*

1. $\Gamma_\alpha(x) = x$ *for every* $\alpha \in \mathbb{R}^+$.
2. *If* $\alpha \geq \beta$ *then* $|\Gamma_\alpha(y_1) - \Gamma_\alpha(y_2) - \Gamma_\alpha(z_1) + \Gamma_\alpha(z_2)| \leq |\Gamma_\beta(y_1) - \Gamma_\beta(y_2) - \Gamma_\beta(z_1) + \Gamma_\beta(z_2)|$ *for every* $y_1, y_2, z_1, z_2 \in \mathbb{R}^d$.
3. Γ_1 *is the identity and* $(\Gamma_\alpha)^{-1} = \Gamma_{\alpha^{-1}}$.
4. *For every compact set* $\mathcal{K} \subset C^\alpha([0,T];\mathbb{R}^d)$, $f \in \mathcal{K}$ *and* $\varepsilon > 0$, $\exists \delta > 0$ *such that* $|pq - 1| < \delta$ *implies*
$$\|\Gamma_p \circ \Gamma_q(f) - f\|_\alpha < \sqrt{\varepsilon}, \qquad \forall\, p, q \in \mathbb{R}^+.$$

The simplest example of such a system of contractions is $\Gamma_\alpha(y) = \frac{y}{\alpha}$ centered at $x = 0$. Indeed this is the specific operator used when proving Strassen's Law for Brownian motion. Also note that we only really care about Γ_α for $\alpha > 1$. It is clear that for $\alpha < 1$, the operators Γ_α will not be contraction operators.

Example 8.2.2. *In fact, a linear contraction operator with a transformation will satisfy these conditions. Consider for example $\Gamma_\alpha(y) = \frac{(y-x)}{\alpha} + x$ and naturally, $\Gamma_\alpha(x) = x$. Similarly, for $\alpha \geq \beta$*

$$\begin{aligned}
&\Gamma_\alpha(y_1) - \Gamma_\alpha(y_2) - \Gamma_\alpha(z_1) + \Gamma_\alpha(z_2)\\
&\qquad = \frac{y_1 - x}{\alpha} + x - \frac{y_2 - x}{\alpha} - x - \frac{z_1 - x}{\alpha} - x + \frac{z_2 - x}{\alpha} + x\\
&\qquad \leq \frac{y_1 - y_2 - z_1 + z_2}{\beta} = \Gamma_\beta(y_1) - \Gamma_\beta(y_2) - \Gamma_\beta(z_1) + \Gamma_\beta(z_2)
\end{aligned}$$

Finally, for $|pq - 1| < \delta$ we have

$$\begin{aligned}
\|\Gamma_p \circ \Gamma_q(f) - f\|_\alpha &= \sup_{s,t\in[0,1]} \frac{|\Gamma_p \circ \Gamma_q(f(t)) - f(t) - \Gamma_p \circ \Gamma_q(f(s)) + f(s)|}{|t-s|^\alpha}\\
&= \sup_{s,t\in[0,1]} \frac{\left|\left[\frac{f(t)}{pq} - f(t)\right] - \left[\frac{f(s)}{pq} - f(s)\right]\right|}{|t-s|^\alpha}\\
&= \left|\frac{1}{pq} - 1\right| \sup_{s,t\in[0,1]} \frac{|f(t) - f(s)|}{|t-s|^\alpha} \leq \frac{\delta}{2}\|f\|_\alpha.
\end{aligned}$$

These conditions are slightly stronger than those of [Bal86] and are used in [EN02]. Condition 2. in Definition 8.2.1 needs to be strengthened to allow it to be applied to Hölder norms rather than just supremum norms. Observe that by choosing $y_2 = z_2 = x$, one gets

$$|\Gamma_\alpha(y_1) - \Gamma_\alpha(z_1)| \leq |\Gamma_\beta(y_1) - \Gamma_\beta(z_1)|.$$

This stronger condition still allows for the example of linear contractions up to a transformation. For $s \in \mathbb{R}^+$ define

$$\phi(s) = \sqrt{s \log(\log(s))}.$$

Let $b : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d'}$ be progressively measurable functions such that there is a unique solution to

$$dY_t = b(Y_t, \mathcal{L}_t^Y)dt + \sigma(Y_t, \mathcal{L}_t^Y)dW_t, \qquad Y_0 = x \in \mathbb{R}^d.$$

Definition 8.2.3. *Let $u > 3$. Let $\hat{\sigma}_u : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d'}$ and $\hat{b}_u : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ be such that*

$$\begin{aligned}
\hat{\sigma}_u(y,\mu) &= \phi(u)\nabla\Big[\Gamma_{\phi(u)}\Big]\Big(\Gamma_{\phi(u)^{-1}}(y)\Big)^T \sigma\Big(\Gamma_{\phi(u)^{-1}}(y), \mu \circ \Gamma_{\phi(u)}\Big)\\
\hat{b}_u(y,\mu) &= u\boldsymbol{L}(y,\mu)\Big[\Gamma_{\phi(u)}\Big]\Big(\Gamma_{\phi(u)^{-1}}(y)\Big),
\end{aligned}$$

where for $\tilde{a} = \sigma^T\sigma$ the operator $\boldsymbol{L}(\cdot,\cdot)[\cdot]$ is given as

$$\begin{aligned}
\boldsymbol{L}(y,\mu)\Big[f\Big]\Big(z\Big) =& \sum_{i=1}^d \frac{\partial f}{\partial y_i}\Big(\Gamma_{\phi(u)^{-1}}(z)\Big) b_i\Big(\Gamma_{\phi(u)^{-1}}(y), \mu\circ\Gamma_{\phi(u)}\Big)\\
&+ \frac{1}{2}\sum_{i,j=1}^d \tilde{a}_{i,j}\Big(\Gamma_{\phi(u)^{-1}}(y), \mu\circ\Gamma_{\phi(u)}\Big)\frac{\partial^2 f}{\partial y_i \partial y_j}\Big(\Gamma_{\phi(u)^{-1}}(z)\Big),
\end{aligned}$$

Assumption 8.2.4. *Throughout we assume that Γ_u is twice differentiable for all $u > 3$ and that $\forall y \in \mathbb{R}^d$, $\forall \mu \in \mathcal{P}_2(\mathbb{R}^d)$ we have for some $\hat{\sigma} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d'}$ and $\hat{b} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$*

$$\lim_{u\to\infty} \hat{\sigma}_u(y,\mu) = \hat{\sigma}(y,\mu) \quad \textit{and} \quad \lim_{u\to\infty} \hat{b}_u(y,\mu) = \hat{b}(y,\mu),$$

where $\hat{\sigma}$ and $\hat{b}$ satisfy Assumption 6.1.1 with the addition that $\hat{\sigma}$ is bounded by constant M.

For $t, u \in \mathbb{R}^+$ define

$$Z_t^u = \Gamma_{\phi(u)}\big(Y_{ut}\big),$$

and recall that since $Y_0 = x$ and $\Gamma_u(x) = x$, by assumption $Z_0^u = x$. We use Itô's formula on Z_t^u by assuming twice differentiability of $\Gamma_{\phi(u)}(\cdot)$.

$$\begin{aligned}
dZ_t^u = d(\Gamma_{\phi(u)}(Y_{ut})) = \nabla\Big[\Gamma_{\phi(u)}\Big]\Big(Y_{ut}\Big)^T dY_{ut}&\\
+ \frac{dY_{ut}^T}{2} H\Big[\Gamma_{\phi(u)}\Big]\Big(Y_{ut}\Big)dY_{ut}&.
\end{aligned}$$

Rewriting $Y_{ut} = \Gamma_{\phi(u)^{-1}}(Z_t^u)$ and substituting in gives

$$\begin{aligned}
dZ_t^u =& u\sum_{i=1}^d \frac{\partial\Gamma_{\Phi(u)}}{\partial y_i}\Big(\Gamma_{\phi(u)^{-1}}(Z_t^u)\Big) b_i\Big(\Gamma_{\phi(u)^{-1}}(Z_t^u), \mathcal{L}_t^{Z^u})\circ\Gamma_{\phi(u)}\Big)dt\\
&+ \frac{u}{2}\sum_{i,j=1}^d \frac{\partial^2\Gamma_{\Phi(u)}}{\partial y_i\partial y_j}\Big(\Gamma_{\phi(u)^{-1}}(Z_t^u)\Big)\sum_{k=1}^{d'}\sigma_{k,i}\sigma_{j,k}\Big(\Gamma_{\phi(u)^{-1}}(Z_t^u), \mathcal{L}_t^{Z^u})\circ\Gamma_{\phi(u)}\Big)dt\\
&+ \sum_{i=1}^d \frac{\partial\Gamma_{\Phi(u)}}{\partial y_i}\Big(\Gamma_{\phi(u)^{-1}}(Z_t^u)\Big)\sum_{k=1}^{d'}\sigma_{i,k}\Big(\Gamma_{\phi(u)^{-1}}(Z_t^u), \mathcal{L}_t^{Z^u})\circ\Gamma_{\phi(u)}\Big)d\mathcal{W}_{ut}^k.
\end{aligned}$$

Next, using that $\mathcal{W}_t^u = \frac{W_{ut}}{\sqrt{u}}$ is a Brownian motion, we can rewrite all of this as the SDE with initial condition $Z_0^u = x$

$$dZ_t^u = \frac{1}{\sqrt{\log\log(u)}}\hat{\sigma}_u\Big(Z_t^u, \mathcal{L}_t^{Z^u})\Big)d\mathcal{W}_t^u + \hat{b}_u\Big(Z_t^u, \mathcal{L}_t^{Z^u})\Big)dt.$$

Under Assumption 8.2.4 and using Theorem 7.2.2 we get

$$-\mathcal{I}(\mathring{A}) \leq \liminf_{u\to\infty} \frac{1}{\log\log(u)} \log \mathbb{P}(Z^u \in A)$$
$$\leq \limsup_{u\to\infty} \frac{1}{\log\log(u)} \log \mathbb{P}(Z^u \in A) \leq -\mathcal{I}(\bar{A}) \tag{8.2.1}$$

for every Borel set A induced by the α-Hölder topology with $\alpha < 1/2$. Recall the definition of the rate function $\mathcal{I}(A) := \inf\big\{\|h\|^2_{\mathcal{H}}/2; h \in \mathcal{H}, \Phi(h). \in A\big\}$ with skeleton process

$$\Phi(h)_t = x + \int_0^t \hat{b}(\Phi(h)_s, \delta_{\Phi(0)_s})ds + \int_0^t \hat{\sigma}(\Phi(h)_s, \delta_{\Phi(0)_s})dh_s.$$

We can now state the main result of this section.

Theorem 8.2.5. *With probability 1, the set of paths $\{Z^u; u > 3\}$ is relatively compact on the Hölder topology and its set of limit points coincides with $K = \Big\{\Phi(h) : \frac{\|h\|^2_{\mathcal{H}}}{2} \leq 1\Big\}$.*

8.3 Auxilliary Lemmas

We first prove some technical Lemmas.

Lemma 8.3.1. *$\forall c > 1$ and $\forall \varepsilon > 0$ there exists a positive integer $j_0(\omega)$ almost surely finite such that $\forall j > j_0$*

$$d_\alpha(Z_{c^j}, K) < \sqrt{\varepsilon}, \quad \textit{where} \quad d_\alpha(x, A) = \inf\Big\{\|x-y\|_\alpha : y \in A\Big\}.$$

Proof. Start by considering the set of α-Hölder continuous paths $C_\varepsilon := \{g; d_\alpha(g, K) \geq \sqrt{\varepsilon}\}$. By definition we have that $\mathcal{I}(C_\varepsilon) > 1$, so there exists a real number $\delta > 0$ such that $\mathcal{I}(C_\varepsilon) > 1 + \delta$. Using the LDP results in (8.2.1), we can rearrange this to get

$$\mathbb{P}\big[Z_{c^j} \in C_\varepsilon\big] \leq \exp\Big(-(1+\delta)\log\log(c^j)\Big) \lesssim \frac{1}{j^{1+\delta}}.$$

Clearly $\sum_{j=1}^\infty \mathbb{P}\big[Z_{c^j} \in C_\varepsilon\big] < \infty$ and by a direct application of Borel-Cantelli we have $\mathbb{P}\big[d_\alpha(Z_{c^j}, K) > \sqrt{\varepsilon} \text{ i.o.}\big] = 0$. ☐

Lemma 8.3.2. *$\forall \varepsilon > 0$ $\exists c_\varepsilon > 1$ such that for $1 < c < c_\varepsilon$ there exists an almost surely finite integer $j_0(\omega)$ such that $\forall j > j_0$, $A_{j,c} \leq \sqrt{\varepsilon}$.*

Proof. For notational convenience define, for $c > 1$ and for every positive integer j, the quantity

$$A_{j,c} = \sup_{c^{j-1}\leq u\leq c^j} \|Z_u - \Gamma_{\phi(u)} \circ \Gamma_{\phi(c^j)^{-1}}(Z_{c^j})\|_\alpha.$$

Start by observing that the set K is relatively compact in the α-topology, so it is bounded. Therefore, by Lemma 8.3.1, we have that $\forall j > j_0$ that $\|Z_{c^j}\|_\alpha < C$. We want to show that

$$\sum_{j\geq 1} \mathbb{P}\Big[A_{c,j} > \sqrt{\varepsilon}\Big] < \infty \quad \text{which is equivalent to} \quad \sum_{j>j_0} \mathbb{P}\Big[A_{c,j} > \sqrt{\varepsilon}, \|Z_{c^j}\|_\alpha < C\Big] < \infty.$$

Considering one of these sets, we see

$$\{A_{j,c} > \sqrt{\varepsilon}, \|Z_{c^j}\|_\infty \leq C\}$$
$$= \Bigg\{ \sup_{c^{j-1}\leq u\leq c^j} \sup_{0\leq s,t\leq 1} \frac{|\Gamma_{\phi(u)}(Y(ut)) - \Gamma_{\phi(u)}(Y(c^j t)) - \Gamma_{\phi(u)}(Y(us) + \Gamma_{\phi(u)}(Y(c^j s))|}{|t-s|^\alpha} > \sqrt{\varepsilon},$$
$$\|Z_{c^j}\|_\alpha < C\Bigg\}.$$

Using Definition 8.2.1, for $u \in [c^{j-1}, c^{j+1}]$

$$\begin{aligned}
&|\Gamma_{\phi(u)}(Y(ut)) - \Gamma_{\phi(u)}(Y(c^j t)) - \Gamma_{\phi(u)}(Y(us) + \Gamma_{\phi(u)}(Y(c^j s))| \\
&\qquad \leq |\Gamma_{\phi(c^{j-1})}(Y(ut)) - \Gamma_{\phi(c^{j-1})}(Y(c^j t)) - \Gamma_{\phi(c^{j-1})}(Y(us) + \Gamma_{\phi(c^{j-1})}(Y(c^j s))|.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\{A_{c,j} > \sqrt{\varepsilon}\} \\
&\subseteq \Bigg\{ \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s,t \leq 1} \frac{1}{|t-s|^\alpha} \Big| \Gamma_{\phi(c^{j-1})}(Y(c^j vt)) - \Gamma_{\phi(c^{j-1})}(Y(c^j t)) \\
&\qquad\qquad - \Gamma_{\phi(c^{j-1})}(Y(c^j vs) + \Gamma_{\phi(c^{j-1})}(Y(c^j s)) \Big| > \sqrt{\varepsilon} \Bigg\} \\
&\subseteq \Bigg\{ \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s,t \leq 1} \frac{|Z_{c^j}(vt) - Z_{c^j}(t) - Z_{c^j}(vs) + Z_{c^j}(s)|}{|t-s|^\alpha} > \frac{\sqrt{\varepsilon}}{2} \Bigg\},
\end{aligned}$$

using that $\exists j$ large enough so that for and $\delta > 0$

$$\frac{\phi(c^{j-1})}{\phi(c^j)} = \frac{1}{\sqrt{c}} \sqrt{\frac{\log\log(c^j)}{\log\log(c^{j-1})}} \leq \frac{1}{\sqrt{c}}(1-\delta),$$

and choosing c small enough we can make $\Gamma_{\frac{\phi(c^{j-1})}{\phi(c^j)}}$ within $\frac{\sqrt{\varepsilon}}{2}$ of the identity operator using properties from Definition 8.2.1. Therefore

$$\begin{aligned}
&\{A_{c,j} > \sqrt{\varepsilon}, \|Z_{c^j}\|_\infty \leq C\} \\
&\subseteq \Bigg\{ \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s,t \leq 1} \frac{|Z_{c^j}(vt) - Z_{c^j}(t) - Z_{c^j}(vs) + Z_{c^j}(t)|}{|t-s|^\alpha} > \frac{\sqrt{\varepsilon}}{2}, \|Z_{c^j}\|_\alpha \leq C \Bigg\} \\
&\subseteq \{Z_{c^j} \in B_\varepsilon\},
\end{aligned}$$

where the set B_ε is given by

$$B_\varepsilon = \Bigg\{ g \in C^\alpha([0,1];\mathbb{R}^d) : \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s,t \leq 1} \frac{|g(vt) - g(t) - g(vs) + g(t)|}{|t-s|^\alpha} > \frac{\sqrt{\varepsilon}}{2}, \|g\|_\alpha \leq C \Bigg\},$$

as we would expect. Let $h \in \mathcal{H}$ so $\|h\|_{\mathcal{H}} < \infty$ such that $\Phi(h) \in B_\varepsilon$, then

$$\begin{aligned}
\frac{\sqrt{\varepsilon}}{2}|t-s|^\alpha &\leq \Big| [\Phi(h)_t - \Phi(h)_{vt}] - [\Phi(h)_s - \Phi(h)_{sv}] \Big| \\
&\leq \Big| \int_{(vt)\vee s}^t d\Phi(h)_r - \int_{vs}^{s\wedge(tv)} d\Phi(h)_r \Big|,
\end{aligned} \tag{8.3.1}$$

for at least some choice of $v \in [\frac{1}{c}, 1]$ and $t, s \in [0,1]$.

We know that a solution to the ODE $\Phi(h)$ exists uniquely and has finite supremum. Therefore we can easily conclude that there exists constants M_1 and M_2 such that

$$\begin{aligned}
\Big| \int_s^t d\Phi(h)_r \Big| &\leq \Big| \int_s^t b(\Phi(h)_r, \delta_{\Phi(h)_r}) dr + \int_s^t \sigma(\Phi(h)_r, \delta_{\Phi(h)_r}) dh_r \Big| \\
&\leq M_1 \|h\|_{\mathcal{H}} \sqrt{|t-s|} + M_2 |t-s|.
\end{aligned}$$

It follows from (8.3.1) that

$$\|h\|_{\mathcal{H}} \geq \frac{\frac{\sqrt{\varepsilon}}{2}|t-s|^{\alpha} - M_2\Big(\Big|t - s \vee (tv)\Big| + \Big|s \wedge (tv) - (sv)\Big|\Big)}{M_1\Big(\Big|t - s \vee (tv)\Big|^{\frac{1}{2}} + \Big|s \wedge (tv) - (sv)\Big|^{\frac{1}{2}}\Big)}.$$

Let us consider first the case where $s < (tv)$.

$$\begin{aligned}
\|h\|_{\mathcal{H}} &\geq \frac{\frac{\sqrt{\varepsilon}}{2}\Big|t-s\Big|^{\alpha} - M_2\Big|(t+s)(1-v)\Big|}{M_1\Big|(\sqrt{t}+\sqrt{s})\sqrt{1-v}\Big|} \\
&\geq \frac{\sqrt{\varepsilon}|1-\frac{1}{c}|^{\alpha}}{4M_1|\sqrt{1-\frac{1}{c}}|} - \frac{M_2}{M_1}\sqrt{1-\tfrac{1}{c}},
\end{aligned}$$

so for c small enough we have $\|h\|_{\mathcal{H}} \geq 1+\delta$ for any choice of $\delta > 0$.

Secondly, consider the case where $s > (tv)$

$$\begin{aligned}
\|h\|_{\mathcal{H}} &\geq \frac{\frac{\sqrt{\varepsilon}}{2}\Big|t(1-\frac{s}{t})\Big|^{\alpha} - M_2\Big(\Big|t(1-\frac{s}{t})\Big|(1+v)\Big)}{2M_1\Big(\Big|t(1-\frac{s}{t})\Big|^{\frac{1}{2}}\Big)} \\
&\geq \frac{\sqrt{\varepsilon}|1-\frac{1}{c}|^{\alpha}}{4M_1|\sqrt{1-\frac{1}{c}}|} - \frac{M_2}{M_1}\sqrt{1-\tfrac{1}{c}},
\end{aligned}$$

and taking $c > 1$ small enough as before gives $\|h\|_{\mathcal{H}} \geq 1+2\delta$.

Therefore, using Equation (8.2.1) we can get

$$\begin{aligned}
\mathbb{P}[Z_{c^j} \in B_\varepsilon] &\leq \exp\Big(-(\mathcal{I}(B_\varepsilon)-\delta)\log\log(c^j)\Big) \\
&\leq \exp\Big(-(1+\delta)\log\log(c^j)\Big) \lesssim \frac{1}{j^{1+\delta}},
\end{aligned}$$

and the conclusion of the proof is straightforward by Borel Cantelli. □

8.4 Proof of Theorem 8.2.5

We are now able to prove the main theorem.

Proof of Theorem 8.2.5. The proof is divided into two parts:

Step 1. Relative Compactness. For any $c > 1$, there will exist $j \in \mathbb{N}$ such that $c^{j-1} < u < c^j$

$$d_\alpha(Z_u, K) \leq d_\alpha(Z_{c^j}, K) \tag{8.4.1}$$
$$+ \|\Gamma_{\phi(u)} \circ \Gamma_{\phi(c^j)^{-1}}(Z_{c^j}) - Z_{c^j})\|_\alpha + \|Z_u - \Gamma_{\phi(u)} \circ \Gamma_{\phi(c^j)^{-1}}(Z_{c^j})\|_\alpha, \tag{8.4.2}$$

where j is chosen so that $c^{j-1} \leq u \leq c^j$.

Lemma 8.3.1 with j large enough ensures that (8.4.1) is bounded by $\frac{\sqrt{\varepsilon}}{3}$. From Lemma 8.3.1, we have that Z_{c^j} is bounded, since $\forall\delta > 0$,

$$1 \geq \frac{\phi(u)}{\phi(c^j)} \geq \frac{\phi(c^{j-1})}{\phi(c^j)} \geq \frac{(1-\delta)}{\sqrt{c}},$$

for j large enough. Choosing $1 < c$ small enough, we can use the forth part of Definition 8.2.1 to get that the 1st term in (8.4.2) is less than $\frac{\sqrt{\varepsilon}}{3}$. Lemma 8.3.2 bounds the 2nd term of (8.4.2) by $\frac{\sqrt{\varepsilon}}{3}$.

Therefore, we conclude that the set $\{Z_u : u > 3\}$ is relatively compact (and hence we have convergence in probability).

Step 2. The set of limit points. Let $\Phi(h) \in K$ so that $\frac{\|h\|_{\mathcal{H}}^2}{2} < 1$. Then for $\varepsilon > 0$ and $\beta > 0$, we define the sets

$$E_j = \Big\{ \Big\| \frac{\mathcal{W}_{c^j}(t)}{\sqrt{\log\log(c^j)}} - h \Big\|_\infty \leq \beta \Big\} \quad \text{and} \quad F_j = \Big\{ \Big\| Z_{c^j} - \Phi(h) \Big\|_\alpha \leq \sqrt{\varepsilon} \Big\}.$$

Using Proposition 7.2.3, we have that for j large enough and α small enough that

$$\mathbb{P}[E_j] - \mathbb{P}[F_j] = \mathbb{P}\Big[E_j \cap F_j^c \Big] \leq \exp\Big(-2\log\log(c^j) \Big) \lesssim \frac{1}{j^2}. \tag{8.4.3}$$

Strassen's Law tells us that $\mathbb{P}\Big[\limsup_{j\in\mathbb{N}} E_j \Big] = 1$, see [Str64]. Therefore $\sum_j \mathbb{P}[E_j] = \infty$. However, by Equation (8.4.3) we also have

$$\begin{aligned} \sum_j \Big(\mathbb{P}[E_j] - \mathbb{P}[F_j] \Big) < \infty \quad &\Rightarrow \quad \sum_j \mathbb{P}\Big[F_j \Big] = \infty \\ &\Rightarrow \quad \mathbb{P}\Big[\Big\| Z_{c^j} - \Phi(h) \Big\|_\alpha < \sqrt{\varepsilon} \text{ i.o.} \Big] = 1, \end{aligned}$$

the latter following from Borel-Cantelli.

Finally since $(c^j)_{j\in\mathbb{N}}$ is just a subsequence of $(m)_{m\in\mathbb{N}}$, the result can be extended to the conclusion.

□

8.5 Examples

Following similar examples to those studied in [Bal86], consider the following Mean-field Levy areas:

Example 8.5.1. *Let* $\Gamma_\alpha\Big[(x,y)\Big] = \Big(\frac{x}{\alpha}, \frac{y}{\alpha^2}\Big)$. *Let* $b = 0$ *and let*

$$\sigma\Big((x,y),\mu\Big) = \begin{pmatrix} 1 & 0 \\ 0 & x + \sqrt{\int |\tilde{x}|^2 \mu(d\tilde{x}, d\tilde{y})} \end{pmatrix}$$

Then

$$Z_t^u = \begin{pmatrix} \frac{W_t^{(1)}}{\sqrt{\log(\log(u))}} \\ \int_0^t \frac{W_s^{(1)} + \mathbb{E}\big[|W_s^{(1)}|^2 \big]^{1/2}}{\log(\log(u))} dW_s^{(2)} \end{pmatrix}$$

has limit points

$$\left\{ \begin{pmatrix} h_1 \\ \int_0^t h_1(s) dh_2(s) \end{pmatrix}, h_1, h_2 \in \mathcal{H}, \|h_1\|_{\mathcal{H}}^2 + \|h\|_{\mathcal{H}}^2 \leq 1 \right\}.$$

Typically for stochastic differential equations, we would expect that in the low probability event that we get a large jump that ensures the driving noise does not converge almost surely, the drift term will magnify this effect either by pushing the process towards infinity or 0. Thus we would only expect to observe a meaningful Functional Iterated Logarithm law for stochastic processes where the drift term acts in the same way around 0 as it does around infinity. This can be seen in the Assumption that $\hat{b}_u \to \hat{b}$.

Part III

First Order Calculus

Chapter 9

Malliavin Differentiability

In this chapter, we close an unexpected gap in the literature concerning the Malliavin and Parametric differentiability of stochastic differential equations with drifts that satisfy a superlinear growth condition. We discovered this as part of a project investigating first-order calculus for McKean Vlasov Equations and emphasise that the extension to the mean-field setting is not the real challenge of this material. Conventional methods for proving Malliavin differentiability for SDEs fail in the case where the drift is not Lipschitz as one ends up dealing with potentially non-integrable error terms. The main advantage of the tools that we use in this Chapter, namely *Stochastic Gâteaux Differentiability* and *Ray Absolute continuity*, is that we are able to avoid these potentially non-integrable terms and establish the relevant limits with sharp conditions.

The results of this Chapter can be found published in [IdRS19, Section 2 and 3].

9.1 Introduction

We present the first two classes of SDEs that we will be working with throughout Part III.

9.1.1 Lipschitz and Locally Lipschitz coefficients

Let $(t, \omega, \theta) \in [0,T] \times \Omega \times L^0(\mathcal{F}_0; \mathbb{P}; \mathbb{R}^d)$.

In this paper, we prove differentiability properties of the SDE

$$X_t^\theta(\omega) = \theta + \int_0^t b\Big(s, \omega, X_s^\theta(\omega)\Big) ds + \int_0^t \sigma\Big(s, \omega, X_s^\theta(\omega)\Big) dW_s, \tag{9.1.1}$$

driven by a d'-dimensional Brownian motion W.

Assumption 9.1.1. *Let $p \geq 2$. Let $\theta : \Omega \to \mathbb{R}^d$, $b : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d \times d'}$ be progressively measurable maps and $L > 0$ such that:*

- $\theta \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$.
- *b and σ are integrable in the sense that*

$$\mathbb{E}\Big[\Big(\int_0^T |b(t,\omega,0)| dt\Big)^p\Big],\ \mathbb{E}\Big[\Big(\int_0^T |\sigma(t,\omega,0)|^2 dt\Big)^{\frac{p}{2}}\Big] < \infty. \tag{9.1.2}$$

- *$\exists L$ such that for almost all $(s,\omega) \in [0,T] \times \Omega$ and $\forall x, y \in \mathbb{R}^d$ we have*

$$\big\langle x - y, b(s,\omega,x) - b(s,\omega,y)\big\rangle_{\mathbb{R}^d} \leq L|x-y|^2 \quad \textit{and} \quad |\sigma(s,\omega,x) - \sigma(s,\omega,y)| \leq L|x-y|.$$

- *For $x, y \in \mathbb{R}^d$ such that $|x|, |y| < N$, $\exists L_N > 0$ such that*

$$|b(s,\omega,x) - b(s,\omega,y)| \leq L_N|x-y|,$$

for almost all $(s,\omega)\in[0,T]\times\Omega$*.*

The next result extends results found in the literature to the case of random coefficients. Existence and uniqueness of a solution follow the methods of [Mao08, Theorem 2.3.6]; the case of random coefficients is not addressed there but the general methodology is applicable in the same way with only more care being taken when proving integrability.

Theorem 9.1.2. *Let* $p\geq 2$*. Suppose Assumption 9.1.1 is satisfied. Then there exists a unique solution* $(X_t)_{t\in[0,T]}$ *to the SDE* (9.1.1) *in* $\mathcal{S}^p$ *and*

$$\mathbb{E}\Big[\|X^\theta\|_\infty^p\Big]\lesssim\bigg(\mathbb{E}\Big[|\theta|^p\Big]+\mathbb{E}\Big[\Big(\int_0^T|b(s,\omega,0)|ds\Big)^p\Big]+\mathbb{E}\Big[\Big(\int_0^T\Big|\sigma(s,\omega,0)\Big|^2ds\Big)^{\frac{p}{2}}\Big]\bigg).$$

Moreover, the map $t\mapsto X_t^\theta(\omega)$ *is* $\mathbb{P}$*-a.s. continuous.*

Finally, the solution of the SDE is Stochastically Stable in the sense that for $\forall\xi,\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$*, (with a constant depending on the other parameters but not on* θ *or* ξ*)*

$$\mathbb{E}\Big[\|X^\theta-X^\xi\|_\infty^p\Big]\lesssim\mathbb{E}\Big[|\theta-\xi|^p\Big].$$

Remark 9.1.3 (Issues with integrability and Fubini - Sharp conditions)**.** *The integrability conditions of Assumption 9.1.1 are designed to be sharp. However, they yield processes which can have some problematic properties.*

It is very important to note that we cannot (in general) swap the order of integration at this point! This is a key point in our manuscript. We are not able to assume that the drift term is sufficiently integrable (given (9.1.2)*) and hence the error terms appearing in the proofs of differentiability below will not be assumed to be integrable.*

To emphasize our point consider the following monotone drift function $b(t,\omega,x)=x-x^5$ *and* $\sigma(t,\omega,x)$ *is chosen so that for some* $t'\in[0,T)$

$$\mathbb{E}\Big[\Big|\int_0^T|\sigma(t,\omega,0)|^2dt\Big|^2\Big]<\infty,\quad\mathbb{E}\Big[\Big|\int_0^{t'}|\sigma(t,\omega,0)|^2dt\Big|^{\frac{5}{2}}\Big]=\infty.$$

These satisfy the conditions of Assumption 9.1.1 for $p=4$ *but not for* $p=5$*. We can then argue as follows: for* $t\in[t',T]$

$$\mathbb{E}[\,|X_t|^4]<\infty,\quad\mathbb{E}[\,|X_t|^5]=\infty\quad\textit{and in particular}\quad\mathbb{E}\Big[\int_{t'}^t|X_s|^5ds\Big]=\infty.$$

The existence of finite fourth moments ensures we have finite first moments and hence for $t>t'$

$$\mathbb{E}\Big[\int_0^t\Big(X_s-X_s^5\Big)ds\Big]<\infty\quad\textit{which implies that}\quad\mathbb{E}\Big[\int_{t'}^tX_s^5ds\Big]<\infty.$$

9.1.2 On SDEs with Linear Coefficients

Let $(t,\omega,\theta)\in[0,T]\times\Omega\times L^0(\mathcal{F}_0;\mathbb{P};\mathbb{R}^d)$ and take an SDE of the form

$$\begin{aligned}X_t^\theta(\omega)=\theta&+\int_0^t\Big[B\big(s,\omega\big)X_s^\theta(\omega)+b(s,\omega)\Big]ds\\&+\int_0^t\Big[\Sigma\big(s,\omega\big)X_s(\omega)+\sigma(s,\omega)\Big]dW_s,\end{aligned}\tag{9.1.3}$$

driven by a d'-dimensional Brownian motion W. The derivatives of SDEs of the form (9.1.1) will satisfy linear SDEs of the form (9.1.3).

Assumption 9.1.4. *Let* $p\geq 1$*. Let* $B:[0,T]\times\Omega\to\mathbb{R}^{d\times d}$*,* $\Sigma:[0,T]\times\Omega\to\mathbb{R}^{(d\times d')\times d}$*,* $b:[0,T]\times\Omega\to\mathbb{R}^d$ *and* $\sigma:[0,T]\times\Omega\to\mathbb{R}^{d\times d'}$ *be progressively measurable maps such that:*

- $\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$.

- *B, b, Σ and σ are integrable in the sense that $\exists L \geq 0$ such that $\forall x \in \mathbb{R}^d$*

$$x^T B(t,\omega)x < L|x|^2 \quad \mathbb{P}\text{-a.s.}, \qquad \int_0^T \|\Sigma(t,\cdot)\|_{L^\infty}^2 dt < \infty,$$
$$\mathbb{E}\Big[\Big(\int_0^T |b(t,\omega)|dt\Big)^p\Big], \quad \mathbb{E}\Big[\Big(\int_0^T |\sigma(t,\omega)|^2 dt\Big)^{\frac{p}{2}}\Big] < \infty.$$

One advantage of SDEs of the form (9.1.3) is that they have an explicit solution unlike SDEs of the form (9.1.1) where a solution exists but cannot be explicitly stated. Linear SDEs do have Lipschitz coefficients, but their Lipschitz constants are not uniform over $(t,\omega) \in [0,T] \times \Omega$. Therefore, we cannot apply Theorem 9.1.2.

Notice that for Assumption 9.1.4, we do not make any requirement on B being positive definite operator. In fact, we may be interested in cases where $\exists x \in \mathbb{R}^d$ such that $x^T(\int_0^T B(t,\omega)dt)x = -\infty$ with positive probability.

Theorem 9.1.5. *Let $p \geq 1$. Suppose Assumption 9.1.4 is satisfied. Then there exists a unique solution $(X_t)_{t\in[0,T]}$ to the SDE (9.1.3) in $\mathcal{S}^p$ with explicit form*

$$X_t^\theta = \Psi_t \left(\theta + \int_0^t \Psi_s^{-1}\Big[b(s,\omega) - \Big\langle \Sigma(s,\omega), \sigma(s,\omega)\Big\rangle_{\mathbb{R}^{d'}}\Big]ds + \int_0^t \Psi_s^{-1}\sigma(s,\omega)dW_s\right),$$

where $\Psi : [0,T] \times \Omega \to \mathbb{R}^{d\times d}$ can be written as

$$\Psi(t) = I_d \exp\left(\int_0^t \Big[B(s,\omega) - \frac{\langle \Sigma(s,\omega), \Sigma(s,\omega)\rangle_{\mathbb{R}^{d'}}}{2}\Big]ds + \int_0^t \Sigma(s,\omega)dW_s\right), \tag{9.1.4}$$

and

$$\mathbb{E}\Big[\|X^\theta\|_\infty\Big] \lesssim \left(\mathbb{E}\Big[|\theta|^p\Big] + \mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big] + \mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big]\right).$$

Moreover, the map $t \mapsto X_t(\omega)$ is $\mathbb{P}$-a.s. continuous.

Finally, the solution X^θ of the equation is Stochastically stable in the sense that $\forall \xi, \theta \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$

$$\mathbb{E}\Big[\|X^\xi - X^\theta\|_\infty^p\Big] \lesssim \mathbb{E}\Big[|\xi - \theta|^p\Big].$$

Proof. An existence and uniqueness proof is found in [IdRS19]. ☐

9.2 A Grönwall inequality

To the best of our knowledge the next result is new and of independent interest. While unsurprising, this is key to the methods of this paper.

Proposition 9.2.1 (Grönwall Inequality for the Topology of Convergence in Probability)**.** *Let $n \in \mathbb{N}$, $A^n : [0,T] \times \Omega \to \mathbb{R}$ be a sequence of adapted stochastic processes such that $\|A^n\|_\infty \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$ (that is conv. in probability: $\forall \varepsilon > 0 \lim_{n\to\infty} \mathbb{P}\big[\|U^n\|_\infty > \varepsilon\big] = 0$).*

Let U^n be the solution of the SDE

$$U_t^n = A_t^n + \int_0^t f\big(U_s^n\big)ds + \int_0^t g\big(U_s^n\big)dW_s, \qquad t \in [0,T]$$

where $f, g : \mathbb{R} \to \mathbb{R}$ are Monotone growth and Lipschitz respectively (see 3rd bullet point of Assumption 9.1.1) and $f(0) = g(0) = 0$.

Then $\|U^n\|_\infty \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$.

Notice that since we do not have finite second moments of $\|A^n\|_\infty$, the result cannot be proved using a mean square type argument.

Proof. Fix $\delta > 0$ and let $n \in \mathbb{N}$. We have for any choice of $\eta > 0$ that

$$\mathbb{P}\Big[\|U^n\|_\infty > \delta\Big] \leq \mathbb{P}\Big[\|U^n\|_\infty > \delta, \|A^n\|_\infty \leq \eta\Big] + \mathbb{P}\Big[\|A^n\|_\infty > \eta\Big].$$

We already have that $\lim_{n\to\infty} \mathbb{P}\big[\|A^n\|_\infty > \eta\big] = 0$ for any choice of $\eta > 0$ by assumption. Define the sequence of stopping times $\tau_n = \inf\{t' > 0 : |A^n_{t'}| > \eta\}$, $n \in \mathbb{N}$.

Firstly, we show that $\lim_{n\to\infty} \tau_n \geq T$ almost surely. Suppose this was not the case. Then $\exists \Omega' \subset \Omega$ with $\mathbb{P}(\Omega') > 0$ and $\forall \omega \in \Omega'$ $\exists n_k(\omega)$ an increasing subsequence of integers such that $\tau_{n_k}(\omega) < T$ for all $k \in \mathbb{N}$. Then $\forall \omega \in \Omega'$, $\|A^{n_k}\|_\infty(\omega) > \eta$ for all $k \in \mathbb{N}$. But that implies that for any $k \in \mathbb{N}$ we have

$$\Omega' \subset \{\omega \in \Omega; \|A^{n_k}\|_\infty(\omega) > \eta\} \quad \text{and hence that} \quad \mathbb{P}\big[\|A^{n_k}\|_\infty > \eta\big] > \mathbb{P}[\Omega'].$$

The latter contradicts the assumption that $\|A^{n_k}\|_\infty$ converges to 0 in probability. So any such set Ω' must have measure 0 and we conclude $\lim_{n\to\infty} \tau_n > T$ almost surely.

The SDE for U^n_t is well defined for $t \in [0, \tau_n]$. Outside of this interval, A^n may not be integrable so we may not be able construct a solution. However $\forall \omega \in \Omega$ such that $\|A^n\|_\infty(\omega) \leq \eta$ we have that $\tau_n(\omega) > T$. Therefore

$$\mathbb{P}\Big[\|U^n_\cdot\|_\infty > \delta, \|A^n\|_\infty \leq \eta\Big] = \mathbb{P}\Big[\|U^n_{\cdot\wedge\tau_n}\|_\infty > \delta, \|A^n\|_\infty \leq \eta\Big],$$

because the process $U^n_\cdot$ and the stopped process $U^n_{\cdot\wedge\tau_n}$ are $\mathbb{P}$-almost surely equal when one restricts to the event where $\|A^n\|_\infty \leq \eta$.

As we know that the solution $U^n_{t\wedge\tau_n}$ will exist and make sense, it serves to introduce this stopping time. Thus we get

$$\begin{aligned}\mathbb{P}\Big[\|U^n\|_\infty > \delta\Big] &\leq \mathbb{P}\Big[\|U^n(\cdot\wedge\tau_n)\|_\infty > \delta, \|A^n\|_\infty \leq \eta\Big] + \mathbb{P}\Big[\|A^n\|_\infty > \eta\Big] \\ &\leq \mathbb{P}\Big[\|U^n_{\cdot\wedge\tau_n}\|_\infty > \delta\Big] + \mathbb{P}\Big[\|A^n\|_\infty > \eta\Big].\end{aligned}$$

Now we consider the SDE for $U^n_{t\wedge\tau_n}$. The stopping time prevents the term $A^n_{t\wedge\tau_n}$ from getting any larger that η and ensures that the stochastic integral is a local martingale. Appealing to Theorem 9.1.2 yields existence/uniqueness of the solution and moment bounds.

$$\mathbb{E}\Big[\|U^n_{\cdot\wedge\tau_n}\|^2_\infty\Big] < \eta^2 e^C \quad \text{and therefore} \quad \mathbb{P}\Big[\|U^n\|_\infty > \delta\Big] \leq \frac{\eta^2 e^C}{\delta^2} + \mathbb{P}\Big[\|A^n\|_\infty > \eta\Big].$$

Choose η such that $\eta^2 e^C/\delta^2 < \varepsilon'/2$. Then find $N \in \mathbb{N}$ such that $\forall n \geq N$ $\mathbb{P}\big[\|A^n\|_\infty > \eta\big] < \varepsilon'/2$. This concludes the proof. ☐

9.3 Malliavin Differentiability of SDEs with monotone coefficients

In this section we prove two Malliavin differentiability result for SDEs in the class given by Assumption 9.1.1. We use a less known method using the concepts of *ray absolute continuity* and *stochastic gâteaux differentiability* initiated by [Sug85] and later developed by [MPR17, IMPR16].

For SDEs of the form (9.1.1), the proof of existence and uniqueness of a solution involves a sequence of random variables which converge almost surely to the solution rather than in mean square. Indeed this sequence of random variables does not converge in mean square, unlike in the proof of Existence and Uniqueness for SDEs with Lipschitz coefficients. This means that the classical method from [Nua06, Lemma 1.2.3] cannot be applied; recall further our observation on the role that Proposition 9.2.1 will play here.

9.3.1 Main results and their assumptions

We state the main assumptions and results with the proofs postponed for later sections.

Assumption 9.3.1. *Let $b : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d\times d'}$ satisfy Assumption 9.1.1 for some $p > 2$. Further, suppose*

(i) For almost all $(t,\omega) \in [0,T] \times \Omega$ the functions $\sigma(t,\omega,\cdot)$ and $b(t,\omega,\cdot)$ have spatial partial derivatives in all directions.

(ii) For all $h \in \mathcal{H}$ and $(\varepsilon, x) \in \mathbb{R}^+ \times \mathbb{R}^d$, we have that the maps $\mathbb{R}^+ \times \mathbb{R}^d \to L^0(\Omega)$

$$(\varepsilon, x) \mapsto \int_0^T \Big|\nabla_x \sigma(t, \omega + \varepsilon h, x)\Big|^2 dt \qquad \text{and} \qquad (\varepsilon, x) \mapsto \int_0^T \Big|\nabla_x b(t, \omega + \varepsilon h, x)\Big|^2 dt,$$

are jointly continuous (where convergence in L^0 means convergence in probability).

(iii) $\exists U : [0,T]^2 \times \Omega \to \mathbb{R}^{d\times d'}$ and $V : [0,T]^2 \times \Omega \to \mathbb{R}^{(d\times d')\times d'}$ which satisfy that for $s > r$ $U(s,r,\omega) = V(s,r,\omega) = 0$ and

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^T \Big|U(s,r,\omega)\Big|^2 ds\Big)^{\frac{1}{2}} dr\Big)^p\Big] < \infty \quad \text{and} \quad \mathbb{E}\Big[\Big(\int_0^T \int_0^T \Big|V(s,r,\omega)\Big|^2 dsdr\Big)^{\frac{p}{2}}\Big] < \infty.$$

(iv) b and σ satisfy, as $\varepsilon \to 0$, that $\forall h \in \mathcal{H}$

$$\mathbb{E}\left[\Big(\int_0^T \Big|\frac{b(r, \omega + \varepsilon h, X_r) - b(r,\omega, X_r)}{\varepsilon} - \int_0^r U(s,r,\omega)\dot{h}(s)ds\Big|dr\Big)^2\right] \to 0,$$
$$\mathbb{E}\left[\int_0^T \Big|\frac{\sigma(r, \omega + \varepsilon h, X_r) - \sigma(r,\omega, X_r)}{\varepsilon} - \int_0^r V(s,r,\omega)\dot{h}(s)ds\Big|^2 dr\right] \to 0.$$

In the above condition neither b or σ are assumed to be in $\mathbb{D}^{1,2}$, they are only assumed to be Malliavin differentiable over the sub-manifold on which X (solution to (9.1.1)) takes values on. After our main results we give examples of SDE illustrating the scope of our assumptions.

The convergence conditions on U and V in Assumption 9.3.1(iii) and (iv) could equivalently been stated in terms of a *ray absolute continuity* and *stochastic gâteaux differentiability* criterion instead of *strong stochastic gâteaux differentiability*.

Theorem 9.3.2 (Malliavin Differentiability of Monotone SDEs)**.** *Let $p > 2$. Let Assumption 9.3.1 hold and denote by X the unique solution of the SDE* (9.1.1) *in $\mathcal{S}^p$.*

Then X is Malliavin differentiable, i.e. $X \in \mathbb{D}^{1,p}(\mathcal{S}^p)$ and there exist adapted processes U and V such that the Malliavin derivative satisfies for $0 \le s \le t \le T$

$$\begin{aligned} D_s X_t(\omega) =& \sigma(s,\omega, X_s(\omega)) + \int_s^t U(s,r,\omega)dr + \int_s^t V(s,r,\omega)dW_r \\ &+ \int_s^t \nabla_x b(r,\omega,X_r(\omega)) D_s X_r(\omega) dr + \int_s^t \nabla_x \sigma(r,\omega,X_r(\omega)) D_s X_r(\omega) dW_r, \end{aligned} \tag{9.3.1}$$

and otherwise $D_s X_t = 0$ for $s > t$.

The proof of Theorem 9.3.2 can be found in Section 9.4.

Remark 9.3.3 (Notation)**.** *At the simplest level, we have X is $\mathbb{R}^d$-valued and W is $\mathbb{R}^{d'}$-valued. Therefore b, σ are $\mathbb{R}^d$- and $\mathbb{R}^{d\times d'}$-valued respectively. Hence we have the collection of one-dimensional SDEs*

$$X_t^{(i)}(\omega) = \theta^{(i)} + \int_0^t b^{(i)}(s,\omega,X_s(\omega))ds + \sum_{j=1}^{d'} \int_0^t \sigma^{(i,j)}(s,\omega,X_s(\omega))dW_s^{(j)},$$

where i is an integer between 1 and d.

The Malliavin Derivative D_sX_t is therefore a $\mathbb{R}^{d\times d'}$ valued process and we get the system of equations

$$\begin{aligned}D_s^{(k)}X_t^{(i)}(\omega) =&\sigma^{(i,k)}(s,\omega,X_s(\omega))ds\\&+\int_s^t U^{(i,k)}(s,r,\omega)dr+\sum_{j=1}^{d'}\int_s^t V^{(i,j,k)}(s,r,\omega)dW_r^{(j)}\\&+\int_s^t \Big\langle(\nabla_x b^{(i)})(r,\omega,X_r(\omega)),D_s^{(k)}X_r(\omega)\Big\rangle_{\mathbb{R}^d}dr\\&+\sum_{j=1}^{d'}\int_s^t \Big\langle(\nabla_x \sigma^{(i,j)})(r,\omega,X_r(\omega)),D_s^{(k)}X_r(\omega)\Big\rangle_{\mathbb{R}^d}dW_r^{(j)},\end{aligned}$$

for i an integer between 1 and d and k an integer between 1 and d'.

Remark 9.3.4 (Mollification and non-differentiability of b and σ)**.** *Using classic mollification arguments the assumptions of Theorem 9.3.2 concerning the behaviour of $x\mapsto b(\cdot,\cdot,x)$ and $x\mapsto\sigma(\cdot,\cdot,x)$ can be further weakened. Namely, σ can be assumed to be uniformly Lipschitz as opposed to continuously differentiable and b can be assumed to have left- and right-derivatives not necessarily equal to each other at every point.*

Under these conditions, a canonical mollification argument allows to re-obtain Theorem 9.3.2 where in (9.3.1) *one replaces $\nabla_x b$ and $\nabla_x\sigma$ by two processes corresponding to their generalized derivatives.*

If b and σ are assumed deterministic then one immediately obtains the familiar result.

Corollary 9.3.5 (Deterministic coefficients case)**.** *Suppose that $b:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d\times d'}$ satisfy Assumption 9.1.1. Further, suppose that $x\mapsto b(\cdot,x)$ and $x\mapsto\sigma(\cdot,x)$ are continuously differentiable in their spatial variables (uniformly in t).*

Then X is Malliavin differentiable and $D_sX_t=0$ for $T\geq s>t\geq 0$ while for $0\leq s\leq t\leq T$

$$\begin{aligned}D_sX_t(\omega)=\sigma(s,X_s(\omega))+\int_s^t\nabla_x b(r,X_r(\omega))D_sX_r(\omega)dr\\+\int_s^t\nabla_x\sigma(r,X_r(\omega))D_sX_r(\omega)dW_r.\end{aligned}$$

9.3.2 Stronger but more tangible assumptions

Assumption 9.3.1 is sharp for our construction, nonetheless, it can be slightly strengthened to Assumption 9.3.6 which is much easier to verify.

Assumption 9.3.6. *Let $b:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^{d\times d'}$ satisfy Assumption 9.1.1 for $p>2$. Further, suppose Assumption 9.3.1 (i) and (ii) hold and*

(iii') b and σ are Malliavin differentiable in the sense that

$$\forall x\in\mathbb{R}^d,\ b(\cdot,\cdot,x)\in\mathbb{D}^{1,p}\big(L^1([0,T];\mathbb{R}^d)\big)\quad\textit{and}\quad\sigma(\cdot,\cdot,x)\in\mathbb{D}^{1,p}\big(L^2([0,T];\mathbb{R}^{d\times d'})\big).$$

(iv') The Malliavin derivatives of b and σ are progressively measurable and Lipschitz in their spacial variables i.e. $\exists L>0$ constant such that $\forall(s,t)\in[0,T]^2$ and $x,y\in\mathbb{R}^d$, $\mathbb{P}$-almost surely

$$\begin{aligned}|D_sb(t,\omega,x)-D_sb(t,\omega,y)|\leq L|x-y|,\\|D_s\sigma(t,\omega,x)-D_s\sigma(t,\omega,y)|\leq L|x-y|.\end{aligned}$$

The second main result of the section is the following theorem.

Theorem 9.3.7. *Let $p > 2$. Let Assumption 9.1.1 hold and denote by X the unique solution of the SDE* (9.1.1) *in $\mathcal{S}^p$. Let b and σ satisfy Assumption 9.3.6. Then the conclusion of Theorem 9.3.2 still holds: $X \in \mathbb{D}^{1,p}(\mathcal{S}^p)$ and DX satisfies $D_s X_t = 0$ for $T \geq s > t \geq 0$ while for $0 \leq s \leq t \leq T$*

$$D_s X_t(\omega) = \sigma\big(s, \omega, X_s(\omega)\big) + \int_s^t (D_s b)\big(r, \omega, X_r(\omega)\big) dr + \int_s^t (D_s \sigma)\big(r, \omega, X_r(\omega)\big) dW_r$$
$$+ \int_s^t \nabla_x b\big(r, \omega, X_r(\omega)\big) D_s X_r(\omega) dr + \int_s^t \nabla_x \sigma\big(r, \omega, X_r(\omega)\big) D_s X_r(\omega) dW_r. \qquad (9.3.2)$$

The proof can be found in Section 9.5. We point out that the mollification Remark 9.3.4 applies to this result as well.

It is a well documented fact, see [Nua06, Theorem 2.2.1], that if one has a SDE with deterministic and Lipschitz drift and diffusion coefficients then the Malliavin derivative is the solution of a homogeneous linear SDE. Both the SDE and the Malliavin Derivative have finite moments of all orders. Therefore the solution of the SDE exists in $\mathbb{D}^{1,\infty}$.

We study the case where the coefficients are random. SDEs of this kind do not always have finite moments of all orders, and the same will apply for the Malliavin derivative. In fact, the integrability of the derivative comes directly from the integrability of the Malliavin derivatives of b and σ.

9.3.3 Examples

In this section, we discuss some interesting examples which emphasize the scope and sharpness of the assumptions made.

Example 9.3.8 (Concerning the continuity of $s \mapsto D_s X_\cdot$)**.** *Previous works on Malliavin calculus, see for example* [Nua06]*, treat the solution of this SDE as being continuous in s. While this is true for those examples studied, it is not true in the general case that we study here. We only have that it is square integrable; this example shows that it is not necessary for the derivative to be continuous in s. Take $g \in L^2([0,T])$ be a deterministic discontinuous function (a step function would be adequate) and assume the one dimensional setting. Consider σ of the form*

$$\sigma(t, \omega, x) = x + \int_0^t g_s dW_s \quad \textit{and} \quad b(t, \omega, x) = 0.$$

Hence X_t satisfies $X_t = 1 + \int_0^t \big[X_s + \int_0^s g_r dW_r\big] dW_s$. It can be shown that the explicit solution of this equation is

$$X_t = \exp\Big(W_t - \frac{t}{2}\Big)\Big[1 - \int_0^t \int_0^r \exp\big(\frac{r}{2} - W_r\big) g_u du dr$$
$$+ \int_0^t \int_0^r \exp\big(\frac{r}{2} - W_r\big) g_u dW_u dr\Big].$$

Note that, as expected, X is a continuous process.

The process V, which represents the Malliavin derivative of σ, is

$$V(s, t, \omega) = D_s \sigma(t, \omega, X_t(\omega)) = g_s \mathbb{1}_{(0,t)}(s)$$
$$\Rightarrow \quad \int_s^t V(s, r, \omega) dW_r = g_s W_{s,t}.$$

Clearly, the latter map is not continuous in s. The Malliavin derivative of X solves

$$D_s X_t = X_s + \int_0^s g_r dW_r + g_s W_{s,t} + \int_s^t D_s X_r dW_r.$$

Define $J_s(t) = \exp\left(W_{s,t} - \frac{t-s}{2}\right)$. *Then the Malliavin derivative has the explicit solution*

$$D_s X_t = J_s(t)\Big[X_s + \int_0^s g_r dW_r + g_s\Big(\int_s^t J_s(r)^{-1} dW_r - \int_s^t J_s(r)^{-1} dr\Big)\Big].$$

Since g *is assumed not to be continuous, this will also not be continuous in* s.

We present a case where the coefficients are not Malliavin differentiable in general but are only differentiable on the set where the solution X takes its values. In other words, Assumption 9.3.1 is satisfied but Assumption 9.3.6 is not.

Example 9.3.9 (Malliavin Differentiable on the right manifold). *Let* $d = m = 1$ *for simplicity. Let* $b(t, \omega, x) = -x$ *and*

$$\sigma(t, \omega(t), x) = \begin{cases} (x-1)^2(x+1)^2 & , x \in [-1, 1] \\ \phi(x) \cdot f(\omega_t) & , |x| > 1 \end{cases},$$

where $\phi \in C^\infty$*,* $\phi(x) = 0$ *for* $|x| \leq 1$ *and* $\phi(x) = 1$ *for* $|x| \geq 2$*. The function* f *is any function* $f : \mathbb{R} \to \mathbb{R}$ *which is bounded, continuous but not differentiable and* ω *is the path of the Brownian motion.*

An example of such a function f *could be*

$$f(x) = \begin{cases} W'(x) & , x \in [-1, 1] \\ -2 & , |x| > 1 \end{cases},$$

where $W'(x)$ *is the Weierstrass function. The Weierstrass function is continuous but not differentiable anywhere and satisfies* $W'(-1) = W'(1) = -2$*. The latter implies that* f *is continuous. Hence* $f(\omega(t))$ *will not be Malliavin differentiable but* $\varepsilon \mapsto f(\omega(t) + \varepsilon h(t))$ *will be continuous.*

The derivative of σ *will satisfy*

$$\partial_x \sigma(t, \omega, x) = \begin{cases} 4x(x-1)(x+1) & , x \in [-1, 1] \\ \phi'(x) \cdot f(\omega(t)) & , 1 < x < 2 \\ 0 & , |x| > 2 \end{cases},$$

so since f *is bounded, we conclude that* σ *is Lipschitz* $\forall \omega \in \Omega$ *and differentiable.*

When the initial conditions determine that the process starts inside the interval $[-1, 1]$*, this is a so-called Wright-Fisher process (see* [MSS12]*) and the solution will remain within the interval* $[-1, 1]$ *with probability 1. This is important because the non-Malliavin Differentiability only affects the system when the process exits the* $[-1, 1]$ *interval. The conditions of Assumption 9.3.1 are satisfied but* $\sigma(\cdot, x)$ *is not Malliavin differentiable for all* $x \in \mathbb{R}^d$.

Remark 9.3.10 (The square-integrability case). *In* [MPR17]*, it is proved that one does not require the ray absolute continuity condition if one can prove a* Strong Stochastic Gâteaux Differentiability *condition, see Theorem 3.3.2 and Equation* (3.3.1)*. However, in* [IMPR16]*, the authors provide a random variable* $Z \in \mathbb{D}^{1,2}$ *which is not strong stochastic gâteaux differentiable in the sense that*

$$\mathbb{E}\Big[\Big|\frac{Z(\omega + \varepsilon h) - Z(\omega)}{\varepsilon} - D^h Z\Big|^2\Big] \nrightarrow 0, \quad \textit{as } \varepsilon \to 0.$$

It is however true that for all values $q \in [1, 2)$

$$\mathbb{E}\Big[\Big|\frac{Z(\omega + \varepsilon h) - Z(\omega)}{\varepsilon} - D^h Z\Big|^q\Big] \to 0, \quad \textit{as } \varepsilon \to 0.$$

In our framework, it is necessary to study the square of increments of the process due to the nature of the monotonicity property. Therefore we require that our SDE has finite moment of order p *for some* $p > 2$*. However, in light of the example provided in* [IMPR16]*, we believe (but do not show) that that there exists a case where the solution to an SDE of the form* (9.1.1) *which has finite moments of order up to* $p = 2$ *which is Malliavin Differentiable. Stochastic gâteaux differentiability would* *follow as before, but it was unclear to us how one would prove ray absolute continuity of such a process.*

Remark 9.3.11 (The spatial Lipschitz condition for the Malliavin Derivatives of b and σ)**.** *In Assumption 9.3.6 (iv') we assume that Db and $D\sigma$ are Lipschitz in the spacial variable. We chose this condition because it is easy to verify and strong enough to ensure that $\forall x \in \mathbb{R}^d$*

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s b(t,\omega,X_t)|^2 ds\Big)^{\frac{1}{2}} dt\Big)^p\Big] < \infty,$$
$$\mathbb{E}\Big[\Big(\int_0^T \int_0^t |D_s \sigma(t,\omega,X_t)|^2 ds dt\Big)^{\frac{p}{2}}\Big] < \infty.$$

However, this condition is by no means necessary. One could consider the case where Db is locally Lipschitz in space and satisfies a linear growth condition and equivalently prove Theorem 9.3.7. However, the proof is more involved as it involves a careful interplay using Hölder's inequality between the maximal integrability of X, Db, $D\sigma$ and several other stochastic terms.

9.4 Proofs of Theorem 9.3.2

In what follows, the choice of θ (the initial condition in (9.1.1)) does not affect the Malliavin derivative because θ is $\mathcal{F}_0$-measurable, see Lemma A.2.2. Further, if Y is $\mathcal{F}_t$-adapted then $D_s Y = 0$ for any $t < s$.

9.4.1 Existence and Uniqueness of the Malliavin derivative $D_s X_t$

We start by establishing that (9.3.1) has a unique solution where X solves (9.1.1). At this point, nothing is said about the solution of (9.3.1) being the Malliavin derivative to X solution of (9.1.1), showing it is the subsequent step.

Theorem 9.4.1. *Let $p > 2$. For $(s,t) \in [0,T]^2$, let X be the solution to the SDE* (9.1.1) *under Assumption 9.3.1. Let $(M_s(t))$ be defined by the matrix of $L^2([0,T])$-valued SDEs*

$$\begin{aligned} M_s(t)(\omega) =& \sigma(s,\omega,X_s(\omega)) + \int_s^t U(s,r,\omega)dr + \int_s^t V(s,r,\omega)dW_r \\ &+ \int_s^t \nabla_x b(r,\omega,X_r(\omega))M_s(r)(\omega)dr + \int_s^t \nabla_x \sigma(r,\omega,X_r(\omega))M_s(r)(\omega)dW_r, \end{aligned} \tag{9.4.1}$$

for $s < t$ and $M_s(t) = 0$ for $s > t$.

Then a unique solution exists in $\mathcal{S}^p([0,T];L^2([0,T]))$ for (9.4.1) *and the process M has finite p^{th} moment, namely*

$$\mathbb{E}\Big[\Big(\sup_{t\in[0,T]} \int_0^T |M_s(t)|^2 ds\Big)^{\frac{p}{2}}\Big] < \infty.$$

Observe that Equation (9.4.1) is linear in M, so the sharpness of the integrability is determined by the integrability of U, V and σ (given the assumed behavior of $\nabla_x b$ and $\nabla_x \sigma$). In the trivial case where $U = V = 0$ and $\sigma = 1$ then M has finite moments of all orders.

Proof of Theorem 9.4.1. For brevity, $t \in [0,T]$ and we omit the explicit ω dependency throughout.

Equation (9.4.1) is an infinite dimensional SDE. We see this when we think of the Malliavin Derivative as being an $L^2([0,T])$ valued stochastic process. Therefore, we need to extend results from Section 9.1 to infinite dimensional spaces. Let e_n be an orthonormal basis of the space $L^2([0,T];\mathbb{R}^{d'})$. This is a separable Hilbert space, so without loss of generality we can say the orthonormal basis is countably infinite. Let V_n be the linear span of the set $\{e_1, ..., e_n\}$. Let

$P_n : L^2([0,T];\mathbb{R}^{d'}) \to V_n$ be the canonical projection operators

$$P_n[f](t) = \sum_{k=1}^{n} \langle f, e_k\rangle_{L^2([0,T];\mathbb{R}^{d'})} e_k(t).$$

Then it is clear that $\lim_{n\to\infty} \|P_n[f] - f\|_{L^2([0,T];\mathbb{R}^{d'})} = 0$. For $k \in \mathbb{N}$, consider the sequence of 1-dimensional Linear Stochastic Differential Equations

$$\begin{aligned} M_k(t) = &\int_0^t \sigma(u, X_u) e_k(u) du + \int_0^t \Big(\int_0^r U(u,r) e_k(u) du \Big) dr + \int_0^t \Big(\int_0^r V(u,r) e_k(u) du \Big) dW_r \\ &+ \int_0^t \nabla_x b(r, X_r) M_k(r) dr + \int_0^t \nabla_x \sigma(r, X_r) M_k(r) dW_r. \end{aligned}$$

These equations are of the same form as (9.1.3), hence a unique solution exists for each k by Theorem 9.1.5. Also, observe that the fundamental matrix Ψ will be the same for each choice of $k \in \mathbb{N}$. Ψ will have the explicit solution

$$\Psi(t) = \exp\Big(\int_0^t \nabla b(r, X_r) dr - \frac{1}{2} \int_0^t \Big\langle \nabla\sigma\big(r, X_r\big), \nabla\sigma\big(r, X_r\big) \Big\rangle_{\mathbb{R}^{d'}} dr + \int_0^t \nabla\sigma(r, X_r) dW_r \Big)$$

and M_k has explicit solution

$$\begin{aligned} M_k(t) =& \Psi(t) \Bigg(\int_0^t \sigma\big(u, X_u\big) e_k(u) du + \int_0^t \Psi(r)^{-1} \Big[\int_0^r U\big(u,r\big) e_k(u) du \\ &- \Big\langle \nabla\sigma\big(r, X_r\big), \int_0^r V\big(u,r\big) e_k(u) du \Big\rangle_{\mathbb{R}^{d'}} \Big] dr + \int_0^t \Psi(r)^{-1} \int_0^r V\big(u,r\big) e_k(u) du dW_r \Bigg). \end{aligned}$$

Next define for $0 \le s,t \le T$, $n \in \mathbb{N}$ the process $M_{(n),s}(t) = \sum_{k=1}^n M_k(t) \otimes e_k(s) \mathbb{1}_{[0,t)}(s)$. This process makes sense as the projection space is finite dimensional so we can rewrite it in a finite dimensional vector form. The solution exists in the space $\mathcal{S}^p(L^2([0,T];\mathbb{R}^{d\times d'}))$ and has the explicit solution

$$\begin{aligned} M_{(n),s}(t) =& \sum_{k=1}^n \Psi(t) \Bigg(\int_0^t \sigma\big(u, X_u\big) e_k(u) du + \int_0^t \Psi(r)^{-1} \Big[\int_0^r U\big(u,r\big) e_k(u) du \\ &- \Big\langle \nabla\sigma\big(r, X_r\big), \int_0^r V\big(u,r\big) e_k(u) du \Big\rangle_{\mathbb{R}^{d'}} \Big] dr \\ &+ \int_0^t \Psi(r)^{-1} \Big[\int_0^r V\big(u,r\big) e_k(u) du \Big] dW_r \Bigg) \otimes e_k(s) \mathbb{1}_{[0,t)}(s), \\ =& \Psi(t) \Bigg(P_n\Big[\sigma(\cdot, X_\cdot)\Big](s) + \int_s^t \Psi(r)^{-1} \Big(P_n\Big[U(\cdot, r)\Big](s) \\ &- \Big\langle \nabla\sigma\big(r, X_r\big), P_n\Big[V\big(\cdot, r\big)\Big](s) \Big\rangle_{\mathbb{R}^{d'}} \Big) dr + \int_s^t \Psi(r)^{-1} P_n\Big[V\big(\cdot, r\big)\Big](s) dW_r \Bigg). \end{aligned}$$

This process satisfies the SDE

$$\begin{aligned} M_{(n),s}(t) =& P_n\Big[\sigma(\cdot, X_\cdot)\Big](s) + \int_s^t P_n\Big[U(\cdot, r)\Big](s) dr + \int_s^t P_n\Big[V(\cdot, r)\Big](s) dW_r \\ &+ \int_s^t \nabla_x b(r, X_r) M_{(n),s}(r) dr + \int_s^t \nabla_x \sigma(r, X(r)) M_{(n),s}(r) dW_r. \end{aligned}$$

For $a^{(i,j)} \in L^2([0,T])$ and for $A(u) = (a^{(i,j)}(u))_{i\in\{1,\dots,d\}, j\in\{1,\dots,d'\}}$, define the norm

$$\|A.\| = \Big(\sum_{i=1}^{d} \sum_{j=1}^{d'} \int_0^T |a^{(i,j)}(u)|^2 du \Big)^{1/2} = \Big(\int_0^T |A(u)|^2 du \Big)^{1/2}. \tag{9.4.2}$$

By Itô's formula, we have

$$\begin{aligned}
&\Big| M_{(n),s}(t) - M_{(m),s}(t) \Big|^2 = \sum_{i=1}^{d} \sum_{k=1}^{d'} \Big| M_{(n),s}^{(i,k)}(t) - M_{(m),s}^{(i,k)}(t) \Big|^2, \\
&= \sum_{i,k} \Big| (P_n - P_m) \Big[\sigma(\cdot, X(\cdot)) \Big]^{(i,k)}(s) \Big|^2 \\
&+ 2 \sum_{i,k} \int_s^t \Big(M_{(n),s}^{(i,k)}(r) - M_{(m),s}^{(i,k)}(r) \Big) \cdot (P_n - P_m) \Big[U(\cdot, r) \Big]^{(i,k)}(s) dr \\
&+ 2 \sum_{i,j,k} \int_s^t \Big(M_{(n),s}^{(i,k)}(r) - M_{(m),s}^{(i,k)}(r) \Big) \cdot (P_n - P_m) \Big[V(\cdot, r) \Big]^{(i,j,k)}(s) dW_r^{(j)} \\
&+ 2 \sum_{i,k} \int_s^t \Big(M_{(n),s}^{(i,k)}(r) - M_{(m),s}^{(i,k)}(r) \Big) \cdot \Big\langle \nabla_x b^{(i)}(r, X_r), M_{(n),s}^{(\cdot,k)}(r) - M_{(m),s}^{(\cdot,k)}(r) \Big\rangle dr \\
&+ 2 \sum_{i,j,k} \int_s^t \Big(M_{(n),s}^{(i,k)}(r) - M_{(m),s}^{(i,k)}(r) \Big) \cdot \Big\langle \nabla_x \sigma^{(i,j)}(r, X_r), M_{(n),s}^{(\cdot,k)}(r) - M_{(m),s}^{(\cdot,k)}(r) \Big\rangle dW_r^{(j)} \\
&+ \sum_{i,j,k} \int_s^t \Big| (P_n - P_m) \Big[V(\cdot, r) \Big]^{(i,j,k)}(s) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X_r), M_{(n),s}^{(\cdot,k)}(r) - M_{(m),s}^{(\cdot,k)}(r) \Big|^2 dr.
\end{aligned}$$

Denote $N_s(t) = M_{(n),s}(t) - M_{(m),s}(t)$ and $(P_n - P_m) = Q$ for brevity. Integrating over s and since every term is positive, we can change the order of integration to obtain

$$\begin{aligned}
&\int_0^t \Big| N_s^{(i,k)}(t) \Big|^2 ds = \int_0^t \Big| Q \big[\sigma(\cdot, \omega, X.) \big]^{(i,k)}(s) \Big|^2 ds \\
&\quad + 2 \int_0^t \int_0^r N_s^{(i,k)}(r) \cdot \Big[Q \big[U(\cdot, r) \big]^{(i,k)}(s) + \Big\langle \nabla_x b^{(i)}(r, X_r), N_s^{(\cdot,k)}(r) \Big\rangle \Big] ds dr \\
&\quad + 2 \sum_j \int_0^t \int_0^r N_s^{(i,k)}(r) \cdot \Big[Q \big[V(\cdot, r) \big]^{(i,j,k)}(s) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X_r), N_s^{(\cdot,k)}(r) \Big\rangle \Big] ds dW_r^{(j)} \\
&\quad + \sum_j \int_0^t \int_0^r \Big| Q \big[V(\cdot, r) \big]^{(i,j,k)}(s) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X_r), N_s^{(\cdot,k)}(r) \Big\rangle \Big|^2 ds dr.
\end{aligned}$$

Next, we use Itô's formula for the function $g(x) = \big(\sum_i x^{(i)} \big)^{p/2}$ to get

$$
\begin{aligned}
\|N.(t)\|^p = \Big(\sum_{i,k}\int_0^t |N_s^{(i,k)}(t)|^2 ds\Big)^{\frac{p}{2}} &= \Big(\int_0^t \Big|Q\big[\sigma(\cdot,\omega,X.)\big](s)\Big|^2 ds\Big)^{\frac{p}{2}} \\
&+ p\int_0^t \|N.(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r N_s^{(i,k)}(r)\Big[Q\big[U(\cdot,r)\big]^{(i,k)}(s) \\
&\qquad + \Big\langle \nabla_x b^{(i)}\big(r,X_r\big), N_s^{(\cdot,k)}(r)\Big\rangle ds\Big)dr \qquad (9.4.3)\\
+\frac{p}{2}\int_0^t \|N.(r)\|^{p-2}\sum_{i,j,k}\int_0^r \Big|Q\big[V(\cdot,r)\big]^{i,j,k}(s) \\
&\qquad + \Big\langle \nabla_x \sigma^{(i,j)}(r,X_r), N_s^{(\cdot,k)}(r)\Big\rangle\Big|^2 dsdr \qquad (9.4.4)\\
+p\int_0^t \|N.(r)\|^{p-2}\sum_{i,j,k}\int_0^r N_s^{(i,k)}(r)\Big(Q\big[V(\cdot,r)\big]^{(i,j,k)}(s) \\
&\qquad + \Big\langle \nabla_x \sigma^{(i,j)}\big(r,X_r\big), N_s^{(\cdot,k)}(r)\Big) dsdW_r \qquad (9.4.5)\\
+p(p-2)\int_0^t \|N.(r)\|^{p-4}\sum_{i,j,k}\Big(\int_0^r N_s^{(i,k)}(r)\Big(Q\big[V(\cdot,r)\big]^{(i,j,k)}(s) \\
&\qquad + \Big\langle \nabla_x \sigma^{(i,j)}\big(r,X_r\big), N_s^{(\cdot,k)}(r)\Big) ds\Big)^2 dr. \qquad (9.4.6)
\end{aligned}
$$

We take a supremum over $t \in [0,T]$ then expectations to show that $\mathbb{E}[\|N\|_\infty^2]$ can be made arbitrarily small for $n, m \in \mathbb{N}$ large enough. Let $a \in \mathbb{N}$ be an integer which we will choose later.

Firstly,

$$
\begin{aligned}
(9.4.3) \le p\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r N_s^{(i,k)}(r)Q\big[U(\cdot,r)\big]^{(i,k)}(s)ds\Big)dr\Big] \qquad (9.4.7)\\
+p\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r N_s^{(i,k)}(r)\Big\langle \nabla_x b^{(i)}\big(r,X_r\big), N_s^{(\cdot,k)}(r)\Big\rangle ds\Big)dr\Big]. \qquad (9.4.8)
\end{aligned}
$$

Now we deal with (9.4.7) using Hölder inequality, the norm (9.4.2), then dominate via the supremum norm and move the term outside the integral to merge it with the outer integrand term

$$
\begin{aligned}
(9.4.7) \le & p\mathbb{E}\Big[\|N.\|_\infty^{p-1}\int_0^T \Big(\sum_{i,k}\int_0^r |Q[U(\cdot,r)]^{(i,k)}(s)|^2 ds\Big)^{\frac{1}{2}} dr\Big] \\
\le & \frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{a} + [a(p-1)]^{p-1}\mathbb{E}\Big[\Big(\int_0^T \|Q[U(\cdot,r)](s)\|_2 ds\Big)^p\Big],
\end{aligned}
$$

and

$$
(9.4.8) \le pL\int_0^T \mathbb{E}\Big[\|N.\|_{\infty,r}^p\Big]dr.
$$

using the Monotonicity property of b. Secondly,

$$
\begin{aligned}
\text{(9.4.4)} \leq & p\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-2}\Big(\sum_{i,j,k}\int_0^r |Q[V(\cdot,r)]^{(i,j,k)}|^2 ds\Big)dr\Big] \\
& + p\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-2}\Big(\sum_{i,j,k}\int_0^r \Big\langle \nabla_x\sigma^{(i,j)}(r,X_r), N_s^{(\cdot,k)}(r)\Big\rangle^2 ds\Big)dr\Big] \\
& \leq \frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{a} + 2[a(p-2)]^{\frac{p-2}{2}}\mathbb{E}\Big[\Big(\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big)^{\frac{p}{2}}\Big] + pL^2\int_0^t \mathbb{E}\Big[\|N.\|_{\infty,r}^p\Big]dr,
\end{aligned}
$$

using the boundedness of $\nabla\sigma$. Thirdly, using the Burkholder-Davis-Gundy Inequality

$$
\begin{aligned}
\text{(9.4.5)} \leq & pC_1\mathbb{E}\Big[\Big(\int_0^T \|N.(r)\|^{2p-4}\sum_j\Big(\sum_{i,k}\int_0^r N_s^{(i,k)}(r)\Big[Q[V(\cdot,r)]^{(i,j,k)}(s) \\
& + \Big\langle \nabla_x\sigma^{(i,j)}(r,X_r), N_s^{(\cdot,k)}(r)\Big\rangle\Big]ds\Big)^2 dr\Big)^{\frac{1}{2}}\Big] \\
\leq & \sqrt{2}pC_1\mathbb{E}\Big[\|N.\|_\infty^{p-2}\Big(\int_0^T\Big[\sum_{i,j,k}\int_0^r |N_s^{(i,k)}(r)|\cdot|Q[V(\cdot,r)]^{(i,j,k)}(s)|ds\Big]^2 dr\Big)^{\frac{1}{2}}\Big] \qquad \text{(9.4.9)} \\
& + \sqrt{2}pC_1\mathbb{E}\Big[\|N.\|_\infty^{p-2}\Big(\int_0^T\Big[\sum_{i,j,k}\int_0^r |N_s^{(i,k)}(r)| \\
& \cdot\Big|\Big\langle \nabla_x\sigma^{(i,j)}(r,X_r), N_s^{(\cdot,k)}(r)\Big\rangle\Big|ds\Big]^2 dr\Big)^{\frac{1}{2}}\Big]. \qquad \text{(9.4.10)}
\end{aligned}
$$

As before, we have

$$
\begin{aligned}
\text{(9.4.9)} \leq & \sqrt{2}pC_1\mathbb{E}\Big[\|N.\|_\infty^{p-1}\Big(\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big)^{\frac{1}{2}}\Big] \\
\leq & \frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{a} + (\sqrt{2}C_1)^p[a(p-1)]^{p-1}\mathbb{E}\Big[\Big(\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big)^{\frac{p}{2}}\Big],
\end{aligned}
$$

and put together

$$
\begin{aligned}
\text{(9.4.10)} \leq & \sqrt{2}pC_1L\mathbb{E}\Big[\Big(\int_0^T \|N.(r)\|^{2p}dr\Big)^{\frac{1}{2}}\Big] \leq \sqrt{2}pC_1L\mathbb{E}\Big[\|N.\|_\infty^{\frac{p}{2}}\Big(\int_0^T \|N.(r)\|^p dr\Big)^{\frac{1}{2}}\Big] \\
\leq & \frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{a} + \frac{(pC_1L)^2 a}{2}\int_0^T \mathbb{E}\Big[\|N.\|_{\infty,r}^p\Big]dr.
\end{aligned}
$$

Finally, for

$$
\begin{aligned}
\text{(9.4.6)} \leq & 2p(p-2)\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-4}\Big(\sum_{i,j,k}\int_0^r |N_s^{(i,k)}(r)|\cdot|Q[V(\cdot,r)]^{(i,j,k)}(s)|ds\Big)^2 dr\Big] \\
& \qquad \text{(9.4.11)} \\
& + 2p(p-2)\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-4}\Big(\sum_{i,j,k}\int_0^r |N_s^{(i,k)}(r)| \\
& \cdot\Big|\Big\langle \nabla_x\sigma^{(i,j)}(r,X_r), N_s^{(\cdot,k)}(r)\Big\rangle\Big|ds\Big)^2 dr\Big]. \qquad \text{(9.4.12)}
\end{aligned}
$$

Repeating the same ideas as before, we get

$$\begin{aligned}
\text{(9.4.11)} \leq & 2p(p-2)\mathbb{E}\Big[\|N.\|_\infty^{p-2} \cdot \int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big] \\
\leq & \frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{a} + 2^{\frac{p+2}{2}} \cdot a^{\frac{p-2}{2}} \cdot (p-2)^{p-1}\mathbb{E}\Big[\Big(\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big)^{\frac{p}{2}}\Big].
\end{aligned}$$

and (9.4.12) $\leq 2p(p-2)L^2\mathbb{E}\big[\int_0^T \|N.\|_{\infty,r}^p dr\big]$. Therefore, choosing $a=6$ we conclude

$$\begin{aligned}
\frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{6} \leq & \Bigg(\mathbb{E}\Big[\|Q[\sigma(\cdot,X.)](\cdot)\|_2^p\Big] + \widetilde{C}_1\mathbb{E}\Big[\Big(\int_0^T \|Q[U(\cdot,r)](\cdot)\|_2 dr\Big)^p\Big] \\
& + \widetilde{C}_2\mathbb{E}\Big[\Big(\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big)^{\frac{p}{2}}\Big]\Bigg) + \widetilde{C}_3\int_0^T \mathbb{E}\Big[\|N.\|_{\infty,r}^p\Big] dr.
\end{aligned}$$

By applying Grönwall's inequality we conclude that

$$\begin{aligned}
\mathbb{E}\Big[\sup_{t\in[0,T]} \|M_{(n),\cdot} - M_{(m),\cdot}\|^p\Big] \lesssim & \Bigg(\mathbb{E}\Big[\|(P_n - P_m)\big[\sigma(\cdot, X(\cdot))\big]\|_2^p\Big] \\
& + \mathbb{E}\Big[\Big(\int_0^T \Big\|(P_n - P_m)[U(\cdot,r)]\Big\|_2 dr\Big)^p\Big] + \mathbb{E}\Big[\Big(\int_0^T \Big\|(P_n - P_m)[V(\cdot,r)]\Big\|_2^2 dr\Big)^{\frac{p}{2}}\Big]\Bigg).
\end{aligned}$$

Given that by assumption we already have

$$\mathbb{E}\Big[\|\sigma(\cdot,X.)\|_2^p\Big],\quad \mathbb{E}\Big[\Big(\int_0^T \|U(\cdot,r)\|_2 dr\Big)^p\Big],\quad \mathbb{E}\Big[\Big(\int_0^T \|V(\cdot,r)\|_2^2 dr\Big)^{\frac{p}{2}}\Big] < \infty,$$

we are able to apply the Dominated Convergence Theorem to swap the order of limits and integrals. Taking a limit as m,n go to infinity lets us conclude that the sequence $M_{(n)}$ is Cauchy in $\mathcal{S}^p(L^2([0,T];\mathbb{R}^{d\times d'}))$. This is a Banach space, so a limit must exist which we denote by M',

$$\begin{aligned}
M_s'(t) = \lim_{n\to\infty} \Psi(t)\Bigg(& P_n\Big[\sigma(\cdot,X.)\Big](s) + \int_s^t \Psi(r)^{-1}\Big(P_n\Big[U(\cdot,r)\Big](s) \\
& - \Big\langle \nabla\sigma\big(r,X_r\big), P_n\Big[V\big(\cdot,r\big)\Big](s)\Big\rangle_{\mathbb{R}^{d'}}\Big)dr + \int_s^t \Psi(r)^{-1}P_n\Big[V\big(\cdot,r\big)\Big](s)dW_r\Bigg).
\end{aligned}$$

Now let $g\in L^2([0,T];\mathbb{R}^{d'})$ be chosen arbitrarily. Then we define $M^{g\prime}(\cdot)$ as

$$\begin{aligned}
M^{g\prime}(t) = & \int_0^t M_s'(t)g(s)ds \\
= & \Psi(t)\Bigg(\int_0^t \sigma(s,X_s)g(s)ds + \int_s^t \Psi(r)^{-1}\Big(\int_0^t U(s,r)g_s ds \\
& - \Big\langle \nabla\sigma\big(r,X_r\big), \int_0^t V\big(s,r\big)g_s ds\Big\rangle_{\mathbb{R}^{d'}}\Big)dr + \int_s^t \Psi(r)^{-1}\int_0^t V\big(s,r\big)g_s ds dW_r\Bigg).
\end{aligned}$$

In order to move the limit inside the different integrals, we use the Dominated Convergence Theorem again.

Given an explicit solution, we know $M^{g\prime}$ will satisfy the SDE

$$\begin{aligned}M^{g\prime}(t) = \int_0^t \sigma(s,X_s)g_s ds + \int_0^t \Big(\int_0^r U(s,r)g_s ds\Big)dr + \int_0^t \Big(\int_0^r V(s,r)g_s ds\Big)dW_r\\ + \int_0^t \nabla_x b(r,X_r)M^{g\prime}(r)dr + \int_0^t \nabla_x \sigma(r,X_r)M^{g\prime}(r)dW_r.\end{aligned}$$

Therefore by a duality argument

$$\begin{aligned}M_s'(t) =& \sigma(s,X_s) + \int_0^t U(s,r)dr + \int_0^t V(s,r)dW_r\\ &+ \int_0^t \nabla_x b(r,X_r)M_s'(r)dr + \int_0^t \nabla_x\sigma(r,X_r)M_s'(r)dW_r,\end{aligned}$$

which is the same SDE as (9.4.1).

Next we prove uniqueness. Suppose that there are two solutions to the SDE (9.4.1), M and M'. Denote $M - M' = \tilde{N}$. Then $\tilde{N}$ will satisfy the linear SDE

$$d\tilde{N}_s(t) = \nabla_x b(t,X_t)\tilde{N}_s(t)dt + \nabla_x\sigma(t,X_t)\tilde{N}_s(t)dW_t, \quad \tilde{N}_s(s) = 0.$$

Let $g \in L^2([0,T];\mathbb{R}^{d'})$ be chosen arbitrarily. Define $\tilde{N}^g(t) = \int_0^t \tilde{N}_s(t)g_s ds$. Clearly, this linear SDE will almost surely be equal to 0 independently of the choice of g. Hence $\tilde{N}$ must also be equal to 0. So $M = M'$ and we have proved uniqueness. □

9.4.2 Ray Absolute Continuity of X

We show that the expectation of $\|\big(X.(\omega+\varepsilon h) - X.(\omega)\big)/\varepsilon\|_\infty^2$ has a bound uniform in ε. This relies on having finite p^{th} moments of the random variable $\|X\|_\infty$ for $p > 2$. If we only have finite second moments, this would not be true in general.

The case $p = 2$ is problematic. It is not the case that $Z \in \mathbb{D}^{1,2}$ implies that $\big(Z(\omega+\varepsilon h) - Z(\omega)\big)/\varepsilon$ converges in mean square as $\varepsilon \searrow 0$, see Remark 9.3.10 and [IMPR16] for in-depth discussion. If we were dealing with the sharp case where the solution of the SDE exists in $\mathcal{S}^2$, it would be unreasonable to expect the Malliavin Derivatives of b and σ to satisfy Assumption 9.3.1(iv), which is necessary for the following Proposition. The power p must be greater that 2, as opposed to 1, because the monotonicity condition lends itself to studying the moments of the SDE for moments of greater than or equal to 2 but is a hindrance for the moments of order less than 2 (computations may involve local times).

Lemma 9.4.2. *Suppose a measurable map $f : \Omega \to E$ is Stochastically Gâteaux Differentiable and additionally that for $\delta > 0$*

$$\sup_{\varepsilon\le 1}\mathbb{E}\left[\Big|\frac{f(\omega+\varepsilon h) - f(\omega)}{\varepsilon}\Big|^{1+\delta}\right] < \infty. \tag{9.4.13}$$

Then f is Malliavin Differentiable (and so f is ray absolutely continuous).

Proof. Condition (9.4.13) yields the collection of random variables $\big((f(\omega+\varepsilon h) - f(\omega))/\varepsilon\big)_{\varepsilon\le 1}$ to be uniformly integrable. Stochastic gâteaux differentiability means that this collection of random variables converges in probability to a limit. Since $\delta > 0$, we conclude that the sequence of random variables converges in mean, or equivalently we have strong stochastic gâteaux differentiability. Theorem 3.3.2 shows this is equivalent to Malliavin Differentiability and Theorem 3.2.3 implies we must have ray absolute continuity. □

Proposition 9.4.3. *Let X be solution to the SDE* (9.1.1) *under Assumption 9.3.1. We have*

$$\mathbb{E}\left[\Big\|\frac{X.(\omega+\varepsilon h) - X.(\omega)}{\varepsilon}\Big\|_\infty^2\right] = O(1) \quad \textit{as } \varepsilon \to 0. \tag{9.4.14}$$

After we have proved stochastic gâteaux differentiability (see Theorem 9.4.4), Lemma 9.4.2 and Equation (9.4.14) will imply ray absolute continuity.

Proof. Let $t \in [0,T]$. Using Assumption 9.3.1, we have

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{b(t,\omega+\varepsilon h, X_t(\omega)) - b(t,\omega,X_t(\omega))}{\varepsilon}\Big|dt\Big)^2\Big] \\
&\quad \le 2\mathbb{E}\Bigg[\|h\|_{\mathcal{H}}^2\Big(\int_0^T\Big(\int_0^t |U(s,t,\omega)|^2 ds\Big)^{\frac{1}{2}} dt\Big)^2 \\
&\qquad + \Big(\int_0^T \Big|\frac{b(t,\omega+\varepsilon h, X_t) - b(t,\omega,X_t)}{\varepsilon} - \int_0^t U(s,t,\omega)\dot{h}(s)ds\Big|dt\Big)^2\Bigg] \le O(1),
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(t,\omega+\varepsilon h, X_t(\omega)) - \sigma(t,\omega,X_t(\omega))}{\varepsilon}\Big|^2 ds\Big] \\
&\quad \le 2\mathbb{E}\Bigg[\|h\|_{\mathcal{H}}^2 \int_0^T\int_0^t |V(s,t,\omega)|^2 dsdt \\
&\qquad + \int_0^T \Big|\frac{\sigma(t,\omega+\varepsilon h, X_t) - \sigma(t,\omega,X_t)}{\varepsilon} - \int_0^t V(s,t,\omega)\dot{h}_s ds\Big|^2 dt\Bigg] \le O(1).
\end{aligned}$$

For notational compactness let us introduce $P_t^\varepsilon(\omega) = \big(X_t(\omega+\varepsilon h) - X_t(\omega)\big)/\varepsilon$. We have

$$\begin{aligned}
P_t^\varepsilon(\omega) = &\int_0^t \sigma(s,\omega,X_s(\omega))\dot{h}_s ds \\
&+ \int_0^t \Big(\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))\Big)\dot{h}_s ds \\
&+ \frac{1}{\varepsilon}\int_0^t \Big(b(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - b(s,\omega,X_s(\omega))\Big)ds \\
&+ \frac{1}{\varepsilon}\int_0^t \Big(\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))\Big)dW_s.
\end{aligned}$$

Using Itô's formula for $f(x) = |x|^2$ we have

$$\Big|P_t^\varepsilon(\omega)\Big|^2 = 2\int_0^t \Big\langle P_s^\varepsilon(\omega), \sigma(s,\omega,X_s)(\omega))\dot{h}_s\Big\rangle ds \tag{9.4.15}$$

$$+ 2\int_0^t \Big\langle P_s^\varepsilon(\omega), \Big(\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X_s(\omega))\Big)\dot{h}_s\Big\rangle ds \tag{9.4.16}$$

$$+ 2\int_0^t \Big\langle P_s^\varepsilon(\omega), \Big(\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega,X_s(\omega))\Big)\dot{h}_s\Big\rangle ds \tag{9.4.17}$$

$$+ 2\int_0^t \Big\langle P_s^\varepsilon(\omega), \frac{b(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - b(s,\omega+\varepsilon h, X_s(\omega))}{\varepsilon}\Big\rangle ds \tag{9.4.18}$$

$$+ 2\int_0^t \Big\langle P_s^\varepsilon(\omega), \frac{b(s,\omega+\varepsilon h, X_s(\omega)) - b(s,\omega,X_s(\omega))}{\varepsilon}\Big\rangle ds \tag{9.4.19}$$

$$+ 2\int_0^t \Big\langle P_s^\varepsilon(\omega), \frac{\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X_s(\omega))}{\varepsilon}dW_s\Big\rangle \tag{9.4.20}$$

$$+ 2\int_0^t \Big\langle P_s^\varepsilon(\omega), \frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega,X_s(\omega))}{\varepsilon}dW_s\Big\rangle \tag{9.4.21}$$

$$+ \int_0^t \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))}{\varepsilon}\Big|^2 ds. \tag{9.4.22}$$

We take a supremum over t then expectations. Let n be an integer that we will choose later. By using a combination of Young's Inequality, Cauchy-Schwartz Inequality, Burkholder-Davis-Gundy Inequality and the continuity properties from Assumption 9.3.1 we find the following upper bounds:

$$\begin{aligned}\text{For (9.4.15)} \Rightarrow&\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_s^\varepsilon(\omega), \sigma(s,\omega,X_s(\omega))\dot{h}_s\Big\rangle\Big| ds\Big]\\ \leq&\frac{\mathbb{E}[\|P^\varepsilon\|_\infty^2]}{n} + n\|h\|_{\mathcal{H}}^2\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,X_s(\omega))\Big|^2 ds\Big]\\ \leq&\frac{\mathbb{E}[\|P^\varepsilon\|_\infty^2]}{n} + 2n\|h\|_{\mathcal{H}}^2\Big(L^2\mathbb{E}\Big[\|X\|_\infty^2\Big] + \mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,0)\Big|^2 ds\Big]\Big),\end{aligned}$$

$$\begin{aligned}\text{For (9.4.16)} \Rightarrow&\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_s^\varepsilon(\omega), \Big(\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X_s(\omega))\Big)\dot{h}_s\Big\rangle\Big| ds\Big]\\ \leq&2L\varepsilon\int_0^T \mathbb{E}\Big[\|P^\varepsilon\|_{\infty,s}^2\Big]\cdot|\dot{h}_s|ds,\end{aligned}$$

$$\begin{aligned}\text{For (9.4.17)} \Rightarrow&\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_s^\varepsilon(\omega), \Big(\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega, X_s(\omega))\Big)\dot{h}_s\Big\rangle\Big| ds\Big]\\ \leq&\frac{\mathbb{E}[\|P^\varepsilon\|_\infty^2]}{n} + n\|\varepsilon h\|_{\mathcal{H}}^2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega, X_s(\omega))}{\varepsilon}\Big|^2 ds\Big],\end{aligned}$$

$$\begin{aligned}\text{For (9.4.18)} \Rightarrow&\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_s^\varepsilon(\omega), \frac{b(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - b(s,\omega+\varepsilon h, X_s(\omega))}{\varepsilon}\Big\rangle\Big| ds\Big]\\ \leq&2L\int_0^T \mathbb{E}\Big[\|P^\varepsilon\|_{\infty,s}^2\Big]ds,\end{aligned}$$

$$\begin{aligned}\text{For (9.4.19)} \Rightarrow&\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_s^\varepsilon(\omega), \frac{b(s,\omega+\varepsilon h, X_s(\omega)) - b(s,\omega, X_s(\omega))}{\varepsilon}\Big\rangle\Big| ds\Big]\\ \leq&\frac{\mathbb{E}[\|P^\varepsilon\|_\infty^2]}{n} + n\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{b(s,\omega+\varepsilon h, X_s(\omega)) - b(s,\omega, X_s(\omega))}{\varepsilon}\Big| ds\Big)^2\Big],\end{aligned}$$

$$\begin{aligned}\text{For (9.4.20)} \Rightarrow&\mathbb{E}\Big[\sup_{t\in[0,T]} 2\int_0^t \Big\langle P_s^\varepsilon, \frac{\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X_s(\omega))}{\varepsilon}dW_s\Big\rangle\Big]\\ \leq&2C_1\mathbb{E}\Big[\|P^\varepsilon\|_\infty\Big(\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X_s(\omega))}{\varepsilon}\Big|^2 ds\Big)^{1/2}\Big]\\ \leq&\frac{\mathbb{E}[\|P^\varepsilon\|_\infty^2]}{n} + nC_1^2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X_s(\omega))}{\varepsilon}\Big|^2 ds\Big]\\ \leq&\frac{\mathbb{E}[\|P^\varepsilon\|_\infty^2]}{n} + nC_1^2\int_0^T \mathbb{E}\Big[\|P^\varepsilon\|_{\infty,s}^2\Big]ds,\end{aligned}$$

$$\begin{aligned}\text{For (9.4.21)} \Rightarrow&\mathbb{E}\Big[\sup_{t\in[0,T]} 2\int_0^t \Big\langle P_s^\varepsilon, \frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega, X_s(\omega))}{\varepsilon}dW_s\Big\rangle\Big]\\ \leq&\frac{\mathbb{E}[\|P^\varepsilon\|_\infty^2]}{n} + nC_1\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega, X_s(\omega))}{\varepsilon}\Big|^2 ds\Big],\end{aligned}$$

$$\text{For (9.4.22)} \Rightarrow \mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))}{\varepsilon}\Big|^2 ds\Big]$$

$$\le 2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h,X_s(\omega))}{\varepsilon}\Big|^2 ds\Big] \tag{9.4.23}$$

$$+ 2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega,X_s(\omega))}{\varepsilon}\Big|^2 ds\Big],$$

and finally that (9.4.23) $\le 2L^2 \int_0^T \mathbb{E}\Big[\|P^\varepsilon\|^2_{\infty,s}\Big] ds$.

Combining all these inequalities and choosing $n = 6$, we have

$$\frac{1}{6}\mathbb{E}\Big[\|P^\varepsilon\|^2_\infty\Big] \le \mathbb{E}\Big[\|A^\varepsilon\|^2_\infty\Big] + \bar{C}_1 \int_0^T \mathbb{E}\Big[\|P^\varepsilon\|^2_{\infty,s}\Big] ds,$$

where $\mathbb{E}\big[\|A^\varepsilon\|^2_\infty\big] = O(1)$ as $\varepsilon \to 0$. Grönwall's inequality yields that $\mathbb{E}[\,\|P^\varepsilon\|^2_\infty] = O(1)$ as $\varepsilon \to 0$. □

9.4.3 Stochastic Gateaux Differentiability of X

Next we prove the convergence in probability statement of Definition 3.2.2.

Theorem 9.4.4. *Let X be solution to the SDE* (9.1.1) *under Assumption 9.3.1 and let $h \in \mathcal{H}$. Then we have as $\varepsilon \to 0$*

$$\Big\|\frac{X_\cdot(\omega+\varepsilon h) - X_\cdot(\omega)}{\varepsilon} - \int_0^\cdot M_s(\cdot)(\omega)\dot{h}_s ds\Big\|_\infty \xrightarrow{\mathbb{P}} 0.$$

Hence X satisfies Definition 3.2.2, i.e. is Stochastically Gâteaux differentiable.

Proof. Let $t \in [0,T]$. To make the proof more readable we introduce several shorthand notations M^h, P_ε and Y_ε, to denote increments and its differences, namely, define

$$M^h_t(\omega) := \int_0^t M_s(t)(\omega)\dot{h}(s)ds, \quad P^\varepsilon_t(\omega) := \tfrac{X_t(\omega+\varepsilon h) - X_t(\omega)}{\varepsilon},$$

and $Y^\varepsilon_t(\omega) := P^\varepsilon_t(\omega) - M^h_t(\omega)$. The proof's goal is to show that $\|Y^\varepsilon_\cdot(\omega)\|_\infty \xrightarrow{\mathbb{P}} 0$ as $\varepsilon \searrow 0$.

Methodologically, we write out the SDE for $Y^\varepsilon_t(\omega) = P^\varepsilon_t(\omega) - M^h_t(\omega)$ which we then break into a sequence of terms that are manipulated individually to yield an final inequality amenable to our Grönwall type result for Convergence in Probability of Proposition 9.2.1.

Firstly, we have

$$\begin{aligned} P^\varepsilon_t(\omega) = &\int_0^t \sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h))\dot{h}_s ds \\ &+ \int_0^t \Big[b(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - b(s,\omega,X_s(\omega))\Big] ds \\ &+ \int_0^t \Big[\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))\Big] dW_s. \end{aligned}$$

This would mean we can decompose the SDE for $Y^\varepsilon = P^\varepsilon - M^h$ as

$$\begin{aligned}
Y_t^\varepsilon(\omega) = P_t^\varepsilon(\omega) - M_t^h(\omega) &= \frac{X_t(\omega+\varepsilon h) - X_t(\omega)}{\varepsilon} - M_t^h(\omega)\\
&= \int_0^t \Big[\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))\Big]\dot{h}_s ds \qquad (9.4.24)\\
&+ \int_0^t \Big[\frac{b(s,\omega+\varepsilon h, X_s(\omega)) - b(s,\omega,X_s(\omega))}{\varepsilon} - \int_0^s U(r,s,\omega)\dot{h}_r dr\Big] ds \qquad (9.4.25)\\
&+ \int_0^t \Big[\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega,X_s(\omega))}{\varepsilon} - \int_0^s V(r,s,\omega)\dot{h}_r dr\Big] dW_s \qquad (9.4.26)\\
&+ \int_0^t \Big[\int_0^1 \nabla_x b(s,\omega+\varepsilon h, \Xi_s) d\xi - \nabla_x b(s,\omega,X_s(\omega))\Big] P_s^\varepsilon(\omega) ds \qquad (9.4.27)\\
&+ \int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega+\varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s,\omega,X_s(\omega))\Big] P_s^\varepsilon(\omega) dW_s \qquad (9.4.28)\\
&+ \int_0^t \nabla_x b(s,\omega,X_s(\omega)) Y_s^\varepsilon(\omega) ds + \int_0^t \nabla_x \sigma(s,\omega,X_s(\omega)) Y_s^\varepsilon(\omega) dW_s,
\end{aligned}$$

where $\Xi_\cdot = X_\cdot(\omega) + \xi[X_\cdot(\omega+\varepsilon h) - X_\cdot(\omega)]$.

Then we take $\sup$ over $t \in [0,T]$. Notice that we will not use an Itô type formula on the SDE, but proving convergence for each of the individual terms.

Firstly we consider the mean convergence of (9.4.24),

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T \Big|\Big[\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))\Big]\dot{h}_s\Big| ds\Big)^2\Big]\\
&\le \|h\|_{\mathcal{H}}^2 \cdot \mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \sigma(s,\omega,X_s(\omega))\Big|^2 ds\Big]\\
&\le 2\|h\|_{\mathcal{H}}^2 \Big(\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega,X_s(\omega))\Big|^2 ds\Big]\\
&\quad + 2L^2 T \mathbb{E}\Big[\|X(\omega+\varepsilon h) - X(\omega)\|_\infty^2\Big]\Big)\\
&\le O\Big(\varepsilon^2\Big) + O\Big(\varepsilon^2\Big),
\end{aligned}$$

hence this random variable converges to zero in mean square as $\varepsilon \to 0$.

The term (9.4.25) converges in mean from Assumption 9.3.1 since as $\varepsilon \to 0$

$$\mathbb{E}\Big[\int_0^T \Big|\frac{b(s,\omega+\varepsilon h, X_s(\omega)) - b(s,\omega,X_s(\omega))}{\varepsilon} - \int_0^s U(r,s,\omega)\dot{h}_r dr\Big| ds\Big] \to 0.$$

The term (9.4.26) converges in mean from Assumption 9.3.1, namely as $\varepsilon \to 0$

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} \Big|\int_0^t \Big[\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega,X_s(\omega))}{\varepsilon} - \int_0^s V(r,s,\omega)\dot{h}_r dr\Big] dW_s\Big|\Big]\\
&\le C_1 \mathbb{E}\Big[\Big(\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X_s(\omega)) - \sigma(s,\omega,X_s(\omega))}{\varepsilon} - \int_0^s V(r,s,\omega)\dot{h}_r dr\Big|^2 ds\Big)^{\frac{1}{2}}\Big] \to 0.
\end{aligned}$$

For equation (9.4.27), we are not able to use mean convergence arguments because the terms $\nabla_x b(s,\omega,x)$ have polynomial growth in x and we will not necessarily have enough finite moments to ensure that this term can be dominated. We already have $\lim_{\varepsilon\to 0} \mathbb{E}[\|X(\omega+\varepsilon h) - X(\omega)\|_\infty] = 0$, so clearly we also have convergence in probability. Also by Proposition A.2.4, we have

$$\int_0^T \Big|\nabla_x b(s,\omega+\varepsilon h, X_s(\omega+\varepsilon h)) - \nabla_x b(s,\omega,X_s(\omega))\Big| ds \xrightarrow{\mathbb{P}} 0.$$

for any choice of $x \in \mathbb{R}^d$. Therefore, by continuity of $\nabla_x b$ from Assumption 9.3.1, we get

$$\int_0^T \Big| \int_0^1 \nabla_x b(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x b(s, \omega, X_s(\omega)) \Big| ds \xrightarrow{\mathbb{P}} 0, \quad \text{as } \varepsilon \to 0.$$

Since we also have finite moments of $\|X(\omega + \varepsilon h) - X(\omega)\|_\infty / \varepsilon$ by Proposition 9.4.3, we can conclude that (9.4.27) converges to zero in probability.

For (9.4.28) we know that σ is Lipschitz so we have $\nabla_x \sigma$ is bounded. Hence, we will not have the same integrability issues as with (9.4.27). Therefore, we use convergence in mean. By the Burkholder-Davis-Gundy Inequality and recalling Proposition 9.4.3 we get

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t' \in [0,T]} \Big| \int_0^{t'} \Big(\int_0^1 \nabla_x \sigma(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s, \omega, X_s(\omega)) \Big) \cdot P_s^\varepsilon(\omega) dW_s \Big| \Big] \\
\leq & C_1 \mathbb{E}\Big[\Big(\int_0^T \Big| \int_0^1 \nabla_x \sigma(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s, \omega, X_s(\omega)) \Big|^2 \cdot \Big| P_s^\varepsilon(\omega) \Big|^2 ds \Big)^{\frac{1}{2}} \Big] \\
\leq & C_1 \mathbb{E}\Big[\Big(\int_0^T \Big| \int_0^1 \nabla_x \sigma(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s, \omega, X_s(\omega)) \Big|^2 ds \Big)^{\frac{1}{2}} \cdot \|P^\varepsilon(\omega)\|_\infty \Big] \\
\leq & C_1 \mathbb{E}\Big[\|P^\varepsilon(\omega)\|_\infty^2 \Big]^{\frac{1}{2}} \cdot \mathbb{E}\Big[\int_0^T \Big| \int_0^1 \nabla_x \sigma(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s, \omega, X_s(\omega)) \Big|^2 ds \Big]^{\frac{1}{2}}.
\end{aligned}$$

In the same way as earlier, by continuity of $\nabla_x \sigma$ from Assumption 9.3.1 and Proposition A.2.4 we get

$$\int_0^T \Big| \int_0^1 \nabla_x \sigma(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s, \omega, X_s(\omega)) \Big|^2 ds \xrightarrow{\mathbb{P}} 0.$$

Also, by boundedness of $\nabla_x \sigma$, we have the immediate domination

$$\int_0^T \Big| \int_0^1 \nabla_x \sigma(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s, \omega, X_s(\omega)) \Big|^2 ds \leq 4L^2 T,$$

so we clearly have uniform integrability of all orders. Hence

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big[\int_0^T \Big| \int_0^1 \nabla_x \sigma(s, \omega + \varepsilon h, \Xi_s) d\xi - \nabla_x \sigma(s, \omega, X_s(\omega)) \Big|^2 ds \Big]^{\frac{1}{2}} = 0.$$

Finally, the SDE for the process $Y_t^\varepsilon(\omega)$ can be written in the convenient form

$$Y_t^\varepsilon(\omega) = A^\varepsilon(\omega) + \int_0^t \nabla_x b(s, \omega, X_s(\omega)) Y_s^\varepsilon(\omega) ds + \int_0^t \nabla_x \sigma(s, \omega, X_s(\omega)) Y_s^\varepsilon(\omega) dW_s,$$

where the sequence A^ε is a sequence of random variables which converge to zero in probability. By Proposition 9.2.1 the random variable $\|Y^\varepsilon\|_\infty$ converges in probability to zero as $\varepsilon \to 0$. □

9.4.4 Strong Stochastic Gâteaux Differentiability

We come to the final result of this Section.

Theorem 9.4.5. *Let X be solution to the SDE* (9.1.1) *under Assumption 9.3.1. Then for any $h \in \mathcal{H}$*

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\left\| \frac{X(\omega + \varepsilon h) - X(\omega)}{\varepsilon} - M^h(\omega) \right\|_\infty \right] = 0.$$

Hence X satisfies Equation (3.3.1)*, i.e. is Strong Stochastically Gâteaux differentiable.*

Proof. By Theorem 9.4.4, we have convergence in Probability. Combining this with Proposition

9.4.3 and Theorem 9.4.1, we have

$$\mathbb{E}\Big[\Big\|\frac{X(\omega+\varepsilon h)-X(\omega)}{\varepsilon}\Big\|_\infty^2\Big],\quad \mathbb{E}\Big[\|M^h(\omega)\|_\infty^2\Big]<\infty.$$

Apply Lemma 9.4.2 to conclude. □

Remark 9.4.6. *Although convergence in probability may seem to be rather a weak result relative to the much stronger Almost sure convergence or convergence in mean square, it is actually the case that we now have both. After all, we proved that the sequence of random variables $\big(X_\cdot(\omega+\varepsilon h)-X_\cdot(\omega)\big)/\varepsilon$ have uniform finite p moments over ε and the limit $D^hX_\cdot$ has finite p moments. Therefore, by standard probability theory we have mean square convergence.*

9.4.5 Proof of the Malliavin differentiability result, Theorem 9.3.2

Proof of Theorem 9.3.2. The proof is straightforward and follows from Theorem 9.4.5 and Theorem 3.3.2. Further, the Malliavin Derivative satisfies the SDE (9.3.1) which has a unique solution as proved in Theorem 9.4.1. □

9.5 Proofs of Theorem 9.3.7

In order to prove the Malliavin differentiability (Theorem 9.3.2) under the weakest possible conditions, we only assumed enough properties to ensure convergence of the stochastic gâteaux derivatives. However, the stochastic gâteaux differentiability conditions for b and σ do not require that b and σ are Malliavin differentiable. These conditions need to be checked by the user on a case-by-case basis. Under slightly stronger conditions, but much easier to verify, we present an argument to establish integrability and convergence of b and σ to prove Theorem 9.3.2.

In [GS16], there is a discussion about how much continuity is required for the spacial variable in the Malliavin Derivatives of b and σ in order to prove Malliavin Differentiability of the solution X. The authors prove results similar to those in this paper using much weaker continuity condition, but in doing so assume the integrability of the terms $D_sb(t,\omega,X_t)$ and $D_s\sigma(t,\omega,X_t)$. In our manuscript, we were unable to ensure integrability of b and σ evaluated at X without the Lipschitz (or otherwise tractable assumptions). Weaker continuity conditions would have allowed for examples where $b(t,\omega,X_t(\omega)$ and $\sigma(t,\omega,X_t(\omega))$ were not adequately integrable. Therefore, for easy to check conditions, we work under Assumption 9.3.6 (iii') and (iv') (see Remark 9.3.11).

For simplicity, we introduce Assumption 9.5.1 which contains all of the relevant properties of Assumption 9.3.6 that we require for this section. The function f represents b or σ depending on the choice of m.

Assumption 9.5.1. *Let $m\in\{1,2\}$. Suppose that $f:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ such that*

(i) $\forall x\in\mathbb{R}^d$ $f(\cdot,\cdot,x)\in\mathbb{D}^{1,p}(L^m([0,T];\mathbb{R}^d))$.

(ii) *f is Locally Lipschitz in the spacial variable i.e $\exists L_N>0$ such that $\forall x,y\in\mathbb{R}^d$ such that $|x|,|y|\le N$ and $\forall t\in[0,T]$,*

$$|f(t,\omega,x)-f(t,\omega,y)|\le L_N|x-y|\quad \mathbb{P}\text{-almost surely.}$$

(iii) *Df are Lipschitz in their spatial variables i.e. $\exists L>0$ constant such that $\forall(s,t)\in[0,T]^2$ and $\forall x,y\in\mathbb{R}^d$,*

$$|D_sf(t,\omega,x)-D_sf(t,\omega,y)|\le L|x-y|\quad \mathbb{P}\text{-almost surely.}$$

9.5.1 Integrability and indistinguishability of the Malliavin Derivative

Lemma 9.5.2. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (9.1.1) *under Assumption 9.1.1 and let f satisfy Assumption 9.5.1. Then*

$$\mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^t |D_s f(t,\omega,X_t(\omega))|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] < \infty.$$

Proof. By the definition of $\mathbb{D}^{1,p}(L^m([0,T];\mathbb{R}^d))$ we have for any $t \in [0,T]$

$$\mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] < \infty.$$

Therefore for some constant C (depending on p, m, T, L) we have

$$\begin{aligned}&\mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^t \Big|D_s f(t,\omega,X_t(\omega))\Big|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big]\\ &\qquad\le 2^{\frac{p-m}{m}} C\Bigg(\mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] + \mathbb{E}\Big[\|X\|_\infty^p\Big]\Bigg) < \infty.\end{aligned}$$

□

We have by Assumption 9.5.1 that for every $x \in \mathbb{R}^d$ the random field $f(\cdot,\cdot,x)$ is a Malliavin differentiable process. However, it is not immediate that we have the same for $f(\cdot,\cdot,X_\cdot(\cdot))$. We first prove an indistinguishability property for when we replace x by $X_\cdot(\omega)$.

Lemma 9.5.3. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (9.1.1) *under Assumption 9.1.1. Let f satisfy Assumption 9.5.1 and recall the directional derivative notation introduced previously, $D^h F = \langle DF, h\rangle$ for any choice of $h \in \mathcal{H}$.*

Then, for $h \in \mathcal{H}$ we have, (t,ω)-almost surely that

$$f\Big(t,\omega+\varepsilon h, X_t(\omega)\Big) - f\Big(t,\omega,X_t(\omega)\Big) = \int_0^\varepsilon D^h f(t,\omega+rh,X_t(\omega))dr.$$

Proof. We have that $\forall x \in \mathbb{R}^d$ that $\exists C_x \subset [0,T]\times\Omega$ with $\mathbb{E}[\int_0^T \mathbb{1}_{C_x}(t,\omega)dt] = 0$, dependent on the choice of x, for which $\forall (t,\omega) \in [0,T]\times\Omega\backslash C_x$ that

$$f(t,\omega+\varepsilon h,x) - f(t,\omega,x) = \int_0^\varepsilon D^h f(t,\omega+rh,x)dr. \tag{9.5.1}$$

We wish to prove that we can choose a null set C which is independent of x outside of which the equality holds. To do this, it suffices to prove almost sure continuity with respect to x of both the left and right hand side of (9.5.1).

Almost sure continuity of the left hand side is immediate since f is locally Lipschitz. For the right hand side, we use the Lipschitz properties of the Malliavin derivative. Let r_i be an enumeration of the rationals $\mathbb{Q}^d$. Then we have $\bigcup_i C_{r_i}$ is also a null set since it is the countable union of null sets. Then for $(t,\omega) \in [0,T]\times\Omega\backslash\Big(\bigcup_i C_{r_i}\Big)$ and $\forall x \in \mathbb{Q}^d$ equation (9.5.1) holds. Then by the continuity of f and its Malliavin derivative we conclude that this also holds $\forall x \in \mathbb{R}^d$. □

9.5.2 Strong Stochastic Gâteaux Differentiability

We start by establishing ray absolute continuity.

Lemma 9.5.4. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (9.1.1) *under Assumption*

9.1.1. Let f satisfy Assumption 9.5.1. Then

$$\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{f(t,\omega+\varepsilon h, X_t(\omega)) - f(t,\omega,X_t(\omega))}{\varepsilon}\Big|^m dt\Big)^{\frac{2}{m}}\Big] = O(1), \quad \textit{as } \varepsilon \searrow 0.$$

Proof. Fix $\varepsilon > 0$. By Lemma 9.5.3, for almost all $\omega \in \Omega$ we have that

$$\int_0^T \big|f(t,\omega+\varepsilon h, X_t(\omega)) - f(t,\omega,X_t(\omega))\big|^m dt = \int_0^T \Big|\int_0^\varepsilon D^h f(t,\omega+rh,X_t(\omega))dr\Big|^m dt.$$

Arguing from this, we have with the help of the directional derivative D^h, Jensen and reverse Jensen inequality,

$$\begin{aligned}
&\Big(\int_0^T \big|f(t,\omega+\varepsilon h, X_t(\omega)) - f(t,\omega,X_t(\omega))\big|^m dt\Big)^{\frac{2}{m}} \\
&\quad= \Big(\int_0^T \Big|\int_0^\varepsilon D^h f(t,\omega+rh,X_t(\omega))dr\Big|^m dt\Big)^{\frac{2}{m}} \\
&\quad\le \varepsilon \int_0^\varepsilon \Big(\int_0^T |D^h f(t,\omega+rh,X_t(\omega))|^m dt\Big)^{\frac{2}{m}} dr \\
&\quad\le \varepsilon \|h\|_{\mathcal{H}}^2 \int_0^\varepsilon \Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega+rh,X_t(\omega))|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{2}{m}} dr \\
&\quad\le 2^{\frac{2}{m}} \varepsilon \|h\|_{\mathcal{H}}^2 \Bigg(\int_0^\varepsilon \Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega+rh,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{2}{m}} dr + \varepsilon \|X(\omega)\|_\infty^2 \cdot T^{\frac{2}{m}+1}\Bigg).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathbb{E}\Big[\frac{1}{\varepsilon^2}\Big(\int_0^T \big|f(t,\omega+\varepsilon h, X_t(\omega)) - f(t,\omega,X_t(\omega))\big|^m dt\Big)^{\frac{2}{m}}\Big] \\
&\qquad\le 2^{\frac{2}{m}} \|h\|_{\mathcal{H}}^2 \mathbb{E}\Big[\frac{1}{\varepsilon}\int_0^\varepsilon \Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega+rh,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{2}{m}} dr\Big] \qquad (9.5.2) \\
&\qquad\quad + 2^{\frac{2}{m}} \|h\|_{\mathcal{H}}^2 T^{\frac{2}{m}+1} \mathbb{E}\Big[\|X(\omega)\|_\infty^2\Big].
\end{aligned}$$

We estimate term (9.5.2) as follows and with the help of Proposition A.2.3

$$\begin{aligned}
(9.5.2) \le& 2^{\frac{2}{m}} \|h\|_{\mathcal{H}}^2 \frac{1}{\varepsilon} \int_0^\varepsilon \mathbb{E}\Big[\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} \cdot \mathcal{E}(r\dot{h})(t) dt\Big]^{\frac{2}{m}} dr \\
\le& 2^{\frac{2}{m}} \|h\|_{\mathcal{H}}^2 \mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big]^{\frac{2}{p}} \frac{1}{\varepsilon} \int_0^\varepsilon \mathbb{E}\Big[\|\mathcal{E}(r\dot{h})(\cdot)\|_\infty^{\frac{p}{p-m}}\Big]^{\frac{2(p-m)}{pm}} dr \\
<& O(1),
\end{aligned}$$

with $\mathcal{E}(r\dot{h})$ denoting the stochastic exponential of $r\dot{h}$ as introduced in (A.2.1).

□

Lemma 9.5.5. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (9.1.1) *under Assumption 9.1.1. Let f satisfy Assumption 9.5.1. Then for $h \in \mathcal{H}$ and any $\delta > 0$*

$$\lim_{\varepsilon\to 0} \mathbb{P}\Big[\Big(\int_0^T \Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f\big(t,\omega+rh,X_t(\omega)\big)dr - D^h f\big(t,\omega,X_t)(\omega)\big)\Big|^m dt > \delta\Big] = 0. \qquad (9.5.3)$$

Proof. By Proposition A.2.4, we know that for any $\delta > 0$ that

$$\lim_{\varepsilon\to 0} \mathbb{P}\Big[\int_0^T \Big|D^h f(t,\omega+\varepsilon h, X_t(\omega+\varepsilon h)) - D^h f(t,\omega,X_t(\omega))\Big|^m dt > \delta\Big] = 0. \qquad (9.5.4)$$

Similarly

$$\lim_{\varepsilon\to 0}\mathbb{P}\Big[\|X(\omega+\varepsilon h)-X(\omega)\|_\infty>\delta\Big]=0,$$

so by Lipschitz continuity of Df we also have

$$\lim_{\varepsilon\to 0}\mathbb{P}\Big[\int_0^T\Big|D^h f(t,\omega+\varepsilon h,X_t(\omega+\varepsilon h))-D^h f(t,\omega+\varepsilon h,X_t(\omega))\Big|^m dt>\delta\Big]=0. \tag{9.5.5}$$

Combining Equations (9.5.4) and (9.5.5), we conclude

$$\lim_{\varepsilon\to 0}\mathbb{P}\Big[\int_0^T\Big|D^h f(t,\omega+\varepsilon h,X_t(\omega))-D^h f(t,\omega,X_t(\omega))\Big|^m dt>\delta\Big]=0.$$

Next, using the Fundamental Theorem of Calculus, we also have

$$\lim_{\varepsilon\to 0}\mathbb{P}\Big[\int_0^T\Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f\big(t,\omega+rh,X_t(\omega)\big)dr-D^h f\big(t,\omega,X_t(\omega)\big)\Big|^m dt>\delta\Big]=0.$$

□

The next result establishes the strong stochastic gâteaux differentiability, see Definition 3.3.1.

Lemma 9.5.6. *Let $m\in\{1,2\}$ and $p>2$. Let X be solution to the SDE* (9.1.1) *under Assumption 9.1.1. Let f satisfy Assumption 9.5.1. Then for $h\in\mathcal{H}$*

$$\lim_{\varepsilon\to 0}\mathbb{E}\Big[\Big(\int_0^T\Big|\frac{f(t,\omega+\varepsilon h,X_t(\omega))-f(t,\omega,X_t(\omega))}{\varepsilon}-D^h f(t,\omega,X_t(\omega))\Big|^m dt\Big)^{\frac{1}{m}}\Big]=0.$$

Proof. First, using Lemma 9.5.3, we have $\mathbb{P}$-almost surely that

$$\begin{aligned}\int_0^T&\Big|\frac{f(t,\omega+\varepsilon h,X_t(\omega))-f(t,\omega,X_t(\omega))}{\varepsilon}-D^h f(t,\omega,X_t(\omega))\Big|^m dt\\&=\int_0^T\Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f\big(t,\omega+rh,X_t(\omega)\big)dr-D^h f(t,\omega,X_t(\omega))\Big|^m dt.\end{aligned}$$

By Lemma 9.5.5, both sides converge to 0 in probability (as $\varepsilon\to 0$).

Next, by Lemma 9.5.2 and Lemma 9.5.4, we have uniform L^1 integrability of this collection of random variables since they are bounded in L^2. Convergence in probability and Uniform Integrability imply convergence in mean. □

9.5.3 Proof of Theorem 9.3.7

Proof of Theorem 9.3.7. The difference between Assumptions 9.3.1 and Assumptions 9.3.6 is (iii') and (iv'). Here we verify that b and σ satisfying Assumption 9.3.6 implies Assumptions 9.3.1.

Lemma 9.5.2 implies Assumptions 9.3.1 (iii) is satisfied. Lemma 9.5.6 implies Assumptions 9.3.1 (iv) is satisfied. In this case, the identification U,V with Db and $D\sigma$ respectively is straightforward.This also means that the existence proof in Theorem 9.4.1 holds so a solution to the SDE (9.3.2) must exist. □

Chapter 10

Parametric differentiability

In this section, we study the differentiability properties of solutions of SDEs with respect to the initial condition. For a detailed exploration of the subject of Stochastic flows, see [Kun90]. The main contribution of this section is to prove similar results for SDEs with only locally Lipschitz and monotone coefficients as opposed to previous results which rely on a Lipschitz condition. Similar problems have been studied in [RS17], [Cer01, Chapter 1] and [Zha16].

The results of this Chapter can be found published in [IdRS19, Section 4].

10.1 Gâteaux and Frechét Differentiability of monotone SDEs

We start by recalling the concept of Gâteaux and Frechét Differentiability for abstract Banach Spaces.

Definition 10.1.1 (Gâteaux and Frechét Differentiability)**.** *Let V and W be Banach spaces and let U be an open subset of V. Let $f : U \to W$. The map f is* Gâteaux differentiable *at $x \in U$ in direction $h \in V$ if the limit*

$$\lim_{\varepsilon \to 0} \frac{f(x+\varepsilon h) - f(x)}{\varepsilon} = \frac{d}{d\varepsilon} f(x+\varepsilon h),$$

exists. The limit is called the Gâteaux derivative in direction h.

The map f is said to be Frechét differentiable at $x \in U$ if there exists a bounded linear operator $A : U \to W$ such that

$$\lim_{\|h\|_V \to 0} \frac{\|f(x+h) - f(x) - Ah\|_W}{\|h\|_V} = 0.$$

The linear operator A is called the Frechét derivative of f at x

Let X^θ be the solution of SDE (9.1.1). We next show that the map $\theta \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \mapsto X^\theta \in \mathcal{S}^p([0,T])$ is Frechét differentiable. As we will be differentiating with respect to θ for this section, we emphasize the dependency on θ.

Assumption 10.1.2. *Let $b : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d \times d'}$ satisfy Assumption 9.1.1 for some $p \geq 2$. Further, suppose*

(i) For almost all $(t,\omega) \in [0,T] \times \Omega$ we have the functions $\sigma(t,\omega,\cdot)$ and $b(t,\omega,\cdot)$ have partial derivatives in all directions.

(ii) For all $x \in \mathbb{R}^d$, we have $\mathbb{P}$-almost sure continuity of the maps

$$x \mapsto \int_0^T \Big|\nabla_x \sigma(t,\omega,x)\Big|^2 dt \quad \textit{and} \quad x \mapsto \int_0^T \Big|\nabla_x b(t,\omega,x)\Big|^2 dt.$$

Theorem 10.1.3. *Let $p \geq 2$ and let $1 \leq q < p$. Let X^θ be the solution of SDE* (9.1.1) *under Assumption 10.1.2 in $\mathcal{S}^q$. Then the map $\theta \to X^\theta$ is Gâteaux Differentiable in direction h and the derivative is equal to $F[h]$ the solution of the SDE* (10.1.1)

Further, the operator $F : L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \to S^q([0,T])$ is the Frechét derivative.

Remark 10.1.4. *It is important to note that we were unable to prove Gâteaux Differentiability in the Banach space $\mathcal{S}^p$. Convergence in $\mathcal{S}^p$ would be equivalent to uniform integrability of the random variable*

$$\Big\| \frac{X^{\theta+h} - X^\theta - F[h]}{\|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}} \Big\|_\infty^p,$$

over all possible choices of $h \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$. Unlike in the case where the coefficients are Lipschitz, see [CM18]*, this is not true.*

The proof is given after several intermediary results. The first results relates to Gâteaux differentiability and its properties, we address the Frechét differentiability afterwards. For the proof once one has established Gâteaux differentiability, extending to Frechét differentiability is remarkably easy. Gâteaux differentiability is the weaker condition and is usually considered the easier property to prove.

10.1.1 Existence and Uniqueness for the candidate process

Theorem 10.1.5. *Let $p \geq 2$ and suppose Assumption 10.1.2 holds. Let X^θ be the solution to* (9.1.1)*. Let $h \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$. Then the SDE*

$$F_t[h] = h + \int_0^t \nabla_x b\Big(s, \omega, X_s^\theta(\omega)\Big) F_s[h] ds + \int_0^t \nabla_x \sigma\Big(s, \omega, X_s^\theta(\omega)\Big) F_s[h] dW_s, \tag{10.1.1}$$

has a unique solution in $\mathcal{S}^p([0,T]; \mathbb{R}^d)$.

Proof. This just follows from Theorem 9.1.5. We simply verify that Assumption 9.1.4 holds:

1. $|\nabla_x \sigma| < L$ by the Lipschitz property we have $\mathbb{E}\big[\int_0^T |\nabla_x \sigma(s, \omega, X_s^\theta)|^2 ds\big] < \infty$.
2. From the differentiability and the monotonicity property of b, we have that $\nabla_x b$ is $\mathbb{P}$-almost surely negative semi-definite*. Therefore, for $z \in \mathbb{R}^d$

$$z^T \Big(\int_0^T \nabla_x b(s, \omega, X_s^\theta) ds \Big) z \leq \int_0^T L|z|^2 ds \leq LT|z|^2,$$

Hence, using the moment estimates we conclude that $\mathbb{E}\big[\|F[h]\|_\infty^p\big] \lesssim \|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}^p$. ☐

Unlike with the Malliavin Derivative, the SDE (10.1.1) is not a general linear stochastic differential equation. As b and σ do not have dependency on θ, we do not have extra terms akin to the Malliavin derivatives Db and $D\sigma$. This means that, unlike the Malliavin Derivative, F has finite moments of all orders provided the initial condition has adequate integrability.

Proposition 10.1.6. *Let $p \geq 2$. Suppose Assumption 10.1.2. Let X_θ be the solution to* (9.1.1)*. The operator $F : L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \to \mathcal{S}^p([0,T])$ defined by $h \mapsto F[h]$ the solution of Equation* (10.1.1)*, is a bounded linear operator $\|F[h]\|_{\mathcal{S}^p} \lesssim \|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}$.*

Proof. Firstly, we show that $F[0]_\cdot = 0_d$ a.s. (0_d is the $\mathbb{R}^d$-vector of zeros). Since $F[0]$ is the solution to the SDE

$$\begin{aligned} F_t[0] =& \int_0^t \nabla_x b(s, \omega, X_s^\theta(\omega)) F_s[0] ds + \int_0^t \nabla_x \sigma(s, \omega, X_s^\theta(\omega)) F_s[0] dW_s, \\ F_0[0] =& 0 \end{aligned}$$

*We do not prove this fact; it is straightforward using inner products and the definition of derivative.

and this SDE has a unique solution, we only need to show that $F_\cdot[0] = 0_d$ is a solution. Clearly we have $\mathbb{P}$-almost surely that

$$\int_0^t \nabla_x b(s,\omega,X_s^\theta(\omega)) \cdot 0_d ds = 0 \quad \text{and} \quad \int_0^t \nabla_x \sigma(s,\omega,X_s^\theta(\omega)) \cdot 0_d dW_s = 0,$$

so this is immediate.

Let $\lambda \in \mathbb{R}$. Next we have

$$\begin{aligned}
&F_t[h_1] + \lambda F_t[h_2] \\
&= h_1 + \lambda h_2 + \int_0^t \nabla_x b(s,\omega,X_s^\theta(\omega)) F_s[h_1] ds + \lambda \int_0^t \nabla_x b(s,\omega,X_s^\theta(\omega)) F_s[h_2] ds \\
&\qquad + \int_0^t \nabla_x \sigma(s,\omega,X_s^\theta(\omega)) F_s[h_1] dW_s + \lambda \int_0^t \nabla_x \sigma(s,\omega,X_s^\theta(\omega)) F_s[h_2] dW_s, \\
&\Big(F[h_1] + \lambda F[h_2] \Big)_t \\
&= (h_1 + \lambda h_2) + \int_0^t \nabla_x b(s,\omega,X_s^\theta(\omega)) \Big(F[h_1] + \lambda F[h_2] \Big)_s ds \\
&\qquad + \int_0^t \nabla_x \sigma(s,\omega,X_s^\theta(\omega)) \Big(F[h_1] + \lambda F[h_1] \Big)_s dW_s,
\end{aligned}$$

which is the same as the SDE for $F[h_1 + \lambda h_2]$. Hence, by existence and uniqueness, the two must be equal up to a null set.

For the boundedness, observe that $\|F[h]\|_{\mathcal{S}^p} \lesssim \|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}$ from Theorem 10.1.5. □

10.1.2 Differentiability of $\theta \mapsto X_\theta$

It is immediate to prove the stochastic stability result that $\mathbb{E}\big[\|X_{\theta+h} - X_\theta\|_\infty^p\big]^{1/p} = O\big(\|h\|_{L^p}\big)$ as $\|h\|_{L^p} \to 0$, see Theorem 9.1.2. Hence we have

$$\lim_{\|h\|_{L^p} \to 0} \mathbb{E}\big[\|X_{\theta+h} - X_\theta\|_\infty^p\big] \to 0.$$

Theorem 10.1.7. *Let $p \geq 2$ and $1 \leq q < p$. Let $h \in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$. Suppose we have Assumption 10.1.2, let X_θ be the solution of the SDE* (9.1.1) *and let $F(t)[h]$ be the solution to the SDE* (10.1.1)*. Then we have*

$$\|X_{\theta+h} - X_\theta - F[h]\|_{\mathcal{S}^q} = o\big(\|h\|_{L^p}\big),$$

and therefore $F[h]$ is the Gâteaux derivative of X.

Proof. Let $t \in [0,T]$. Define $\Xi = X^\theta + \xi[X^{\theta+h} - X^\theta]$ and consider

$$\begin{aligned}
&\frac{X_t^{\theta+h} - X_t^\theta - F_t[h]}{\|h\|_{L^p}} = \frac{(\theta+h) - \theta - h}{\|h\|_{L^p}} \\
&+ \int_0^t \Big[\int_0^1 \nabla_x b(s,\omega,\Xi_s) d\xi - \nabla_x b(s,\omega,X_s^\theta) \Big] \cdot \Big[\tfrac{X_s^{\theta+h} - X_s^\theta}{\|h\|_{L^p}} \Big] ds &\quad (10.1.2)\\
&+ \int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega,\Xi_s) d\xi - \nabla_x \sigma(s,\omega,X_s^\theta) \Big] \cdot \Big[\tfrac{X_s^{\theta+h} - X_s^\theta}{\|h\|_{L^p}} \Big] dW_s &\quad (10.1.3)\\
&+ \int_0^t \nabla_x b(s,\omega,X_s^\theta) \Big[\tfrac{X_s^{\theta+h} - X_s^\theta - F_s[h]}{\|h\|_{L^p}} \Big] ds + \int_0^t \nabla_x \sigma(s,\omega,X_s^\theta) \Big[\tfrac{X_s^{\theta+h} - X_s^\theta - F_s[h]}{\|h\|_{L^p}} \Big] dW_s.
\end{aligned}$$

Arguing the same way as in Theorem 9.4.4, we show that Equation (10.1.2) and (10.1.3) converge to zero in probability as $\|h\|_{L^p} \to 0$. Then we apply Proposition 9.2.1 to conclude that

$$\frac{\|X^{\theta+h} - X^\theta - F[h]\|_\infty}{\|h\|_{L^p}} \xrightarrow{\mathbb{P}} 0.$$

Finally, from Theorem 9.1.2 and Theorem 10.1.5 we have that

$$\frac{\mathbb{E}\Big[\|X^{\theta+h} - X^\theta\|_\infty^p\Big]}{\|h\|_{L^p}^p} = O(1), \quad \frac{\mathbb{E}\Big[\|F[h]\|_\infty^p\Big]}{\|h\|_{L^p}^p} = O(1) \qquad \text{as} \quad \|h\|_{L^p} \to 0.$$

Therefore, the random variable $\left\|\frac{X_t^{\theta+h} - X_t^\theta - F_t[h]}{\|h\|_{L^p}}\right\|_\infty^q$ is uniformly integrable and we conclude

$$\left\|\frac{X^{\theta+h} - X^\theta - F[h]}{\|h\|_{L^p}}\right\|_{\mathcal{S}^q} \to 0.$$

□

10.1.3 Proof of the Frechét differentiability theorem

Proof of Theorem 10.1.3. In Proposition 10.1.6 we proved that F is a bounded linear operator and in Theorem 10.1.7 we proved that it satisfies Definition 10.1.1. □

10.2 Classical differentiability of SDEs

For this section, we will be studying the specific case where $\theta = x$ (a constant point in $\mathbb{R}^d$) and our perturbations are all in the constant function directions. Fix $(t, \omega) \in [0, T] \times \Omega$ and consider the map $x \in \mathbb{R}^d \mapsto X_t^x(\omega)$. We will be proving that, with probability 1 and for Lebesgue almost all $t \in [0, T]$, it is a diffeomorphism from $\mathbb{R}^d$ to $\mathbb{R}^d$. For this section, $h \in \mathbb{R}^d$ will represent some deterministic vector in euclidean space. We will be calculating the partial derivatives in direction $h \in \mathbb{R}^d$.

10.2.1 The Jacobian Matrix J

Firstly, we introduce the necessary Assumptions for the Jacobian to exist.

Assumption 10.2.1. *Let $b : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d \times d'}$ satisfy Assumption 9.1.1 for some $p \geq 2$. Further, suppose that $\nabla_x b : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$ and $\nabla_x \sigma : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d \times d' \times d}$ are progressively measurable and that*

(i) For almost all $(t, \omega) \in [0, T] \times \Omega$ we have the functions $\sigma(t, \omega, \cdot)$ and $b(t, \omega, \cdot)$ have partial derivatives in all directions.

(ii) For $x \in \mathbb{R}^d$, we have that the maps $\mathbb{R}^d \to L^0(\Omega)$

$$x \mapsto \int_0^T \Big|\nabla_x \sigma(t, \omega, x)\Big|^2 dt \qquad \text{and} \qquad x \mapsto \int_0^T \Big|\nabla_x b(t, \omega, x)\Big|^2 dt,$$

are continuous (where convergence in L^0 means convergence in probability).

(iii) For almost all $(t, \omega) \in [0, T] \times \Omega$ we have

$$|\nabla_x \sigma(t, \omega, x) - \nabla_x \sigma(t, \omega, y)| \leq L|x - y|.$$

(iv) For $x, y \in \mathbb{R}^d$ such that $|x|, |y| < N$ and for almost all $(s, \omega) \in [0, T] \times \Omega$, $\exists L_N > 0$ such that

$$|\nabla_x b(s, \omega, x) - \nabla_x b(s, \omega, y)| \leq L_N|x - y|.$$

Definition 10.2.2. *Let $p \geq 2$. Let X^x be solution to the SDE* (9.1.1) *under Assumption 10.1.2 and with initial condition $X_0^x = x \in \mathbb{R}^d$. Let I_d be the d-dimensional identity matrix and let $t \in [0,T]$. For $q \geq 1$ and let $J \in \mathcal{S}^q([0,T];\mathbb{R}^{d\times d})$ be the solution of the matrix valued SDE*

$$J_t = I_d + \int_0^t \nabla_x b\Big(s,\omega,X_s^x(\omega)\Big)J_s ds + \int_0^t \nabla_x \sigma\Big(s,\omega,X_s^x(\omega)\Big)J_s dW_s. \tag{10.2.1}$$

Notice that Equation (10.2.1) is the same SDE as (9.1.4). This means the Jacobian has an explicit solution which will be useful in Section 11 below.

Theorem 10.2.3. *Let $p \geq 2$. Let X^x be solution to the SDE* (9.1.1) *under Assumption 10.2.1 and with initial condition $x \in \mathbb{R}^d$. Then the SDE* (10.2.1) *has a unique solution in $\mathcal{S}^p$ and for any choice of $t \in [0,T]$ the map $x \mapsto X_t^x(\omega)$ is differentiable $\mathbb{P}$-almost surely. The derivative is almost surely equal to the solution of the Jacobian Equation, SDE* (10.2.1).

The proof of Theorem 10.2.3 follows from Proposition 10.2.4 and Theorem 10.2.5, see below.

10.2.2 Differentiability of X_x

In the previous section we proved almost sure continuity of $\|X^{x+\varepsilon h} - X^x\|_\infty/\varepsilon$, we need to show that the limit as $\varepsilon \to 0$ is equal to the solution of the Jacobian SDE.

Proposition 10.2.4. *Let $p \geq 2$. Let X^x be solution to the SDE* (9.1.1) *under Assumption 10.2.1 and with initial condition $x \in \mathbb{R}^d$. Then for any choice of $h \in \mathbb{R}^d$, the map*

$$\varepsilon \mapsto \Big\| \frac{X^{x+\varepsilon h}(\omega) - X^x(\omega)}{\varepsilon} \Big\|_\infty,$$

can be extended to when $\varepsilon = 0$ and the extension is almost surely continuous.

Proof. By the Stochastic Stability from Theorem 9.1.2 we have $\mathbb{E}\big[\|X^x - X^y\|_\infty^p\big] \lesssim |x-y|^p$, hence by the Kolmogorov Continuity Criterion we have that the map $\varepsilon \mapsto X^{x+\varepsilon h}$ is almost surely continuous. In fact, one can show α-Hölder continuity for $\alpha < 1$ but not for when $\alpha = 1$ (which would imply Lipschitz Continuity). Therefore, we additionally need to prove almost sure continuity of the map $\varepsilon \mapsto (X^{x+\varepsilon h} - X^x)/\varepsilon$.

Denote for any $t \in [0,T]$ the auxiliary process $K_t^\varepsilon = (X_t^{x+\varepsilon h} - X_t^x)/\varepsilon$. This process satisfies the Linear SDE

$$\begin{aligned} K_t^\varepsilon = h &+ \int_0^t \Big[\int_0^1 \nabla_x b\big(s,\omega,X_s^x + \xi[X_s^{x+\varepsilon h} - X_s^x]\big)d\xi\Big]K_s^\varepsilon ds \\ &+ \int_0^t \Big[\int_0^1 \nabla_x \sigma\big(s,\omega,X_s^x + \xi[X_s^{x+\varepsilon h} - X_s^x]\big)d\xi\Big]K_s^\varepsilon dW_s, \end{aligned}$$

and, introducing the auxiliary process $\Xi^\varepsilon := X^x + \xi[X^{x+\varepsilon h} - X^x]\big)$, we can write the explicit solution of K^ε (as it is the solution a geometric Brownian motion type SDE)

$$K_t^\varepsilon = h \cdot \exp\left(\int_0^t \Big[\int_0^1 \nabla_x b\Big(s,\omega,\Xi_s^\varepsilon\Big)d\xi\Big]ds\right) \cdot \mathcal{E}\left(\int_0^1 \nabla_x \sigma\Big(\cdot,\omega,\Xi_\cdot^\varepsilon\Big)d\xi\right)_t, \tag{10.2.2}$$

where $\mathcal{E}$ is the Doléan-Dade operator introduced in (A.2.1), which for shorthand notation we denote $(K')_t^\varepsilon = \mathcal{E}\big(\int_0^1 \nabla_x\sigma\big(\cdot,\omega,\Xi_\cdot^\varepsilon\big)d\xi\big)_t$.

We now analyze the behaviour of differences of increments of $(K')^\varepsilon$ in ε parameter. Take $\delta > 0$, using the properties of the Doléan-Dade exponential, we have

$$\begin{aligned} (K')_t^\varepsilon - (K')_t^\delta = &\int_0^t \Big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\varepsilon\Big) - \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\Big](K')_s^\varepsilon dW_s \\ &+ \int_0^t \Big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\Big] \cdot \Big[(K')_s^\varepsilon - (K')_s^\delta\Big]dW_s. \end{aligned}$$

Applying Itô's formula for $f(x) = |x|^p$ and denoting $L = (K')^\varepsilon - (K')^\delta$ we get

$$|L_t|^p = p\int_0^t |L_s|^{p-2} L_s^T \cdot \Big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\varepsilon\Big) - \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\Big](K')_s^\varepsilon dW_s \tag{10.2.3}$$

$$+ p\int_0^t |L_s|^{p-2} L_s^T \cdot \Big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\Big]\cdot L_s dW_s \tag{10.2.4}$$

$$+ \frac{p}{2}\int_0^t |L_s|^{p-2}\Big|\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\varepsilon\Big) - \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\big](K')_s^\varepsilon\Big|^2 ds \tag{10.2.5}$$

$$+ \frac{p}{2}\int_0^t |L_s|^{p-2}\Big|\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\big]\cdot L_s\Big|^2 ds \tag{10.2.6}$$

$$+ \frac{p(p-2)}{2}\int_0^t |L_s|^{p-4}\Big|L_s^T\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\varepsilon\Big) - \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\big](K')_s^\varepsilon\Big|^2 ds \tag{10.2.7}$$

$$+ \frac{p(p-2)}{2}\int_0^t |L_s|^{p-4}\Big|L_s^T\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\big]\cdot L_s\Big|^2 ds. \tag{10.2.8}$$

Next, take a supremum over time then expectations. Using the methods that have already been explored in detail for the proof of Theorem 9.4.1, we know that the terms from lines (10.2.4), (10.2.6) and (10.2.8) will all yield terms of the form $\lesssim \int_0^T \mathbb{E}\big[\|L\|_{\infty,t}^p\big]dt$ which will be accounted for with the Grönwall inequality.

Firstly, following the same methods for Theorem 9.4.1 and using the additional Assumption 10.2.1(iii), for

$$\begin{aligned}
\text{for (10.2.3)} \Rightarrow\ & p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |L_s|^{p-2}\cdot\Big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\varepsilon\Big) - \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\Big](K')_s^\varepsilon dW_s\Big],\\
&\le pC_1\mathbb{E}\Big[\Big(\int_0^T |L_s|^{2p-4}\cdot\Big|\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\varepsilon\Big) - \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\big](K')_s^\varepsilon\Big|^2 ds\Big)^{\frac{1}{2}}\Big],\\
&\le \frac{\mathbb{E}\Big[\|L\|_\infty^p\Big]}{n} + C_1^p\big[n(p-1)\big]^{p-1}\mathbb{E}\Big[\Big(TL\|X^{x+\varepsilon h} - X^{x+\delta h}\|_\infty^2\cdot\|(K')^\varepsilon\|_\infty^2\Big)^{\frac{p}{2}}\Big],\\
&\le \frac{\mathbb{E}\Big[\|L\|_\infty^p\Big]}{n} + C_1^p\big[n(p-1)\big]^{p-1}(TL)^{\frac{p}{2}}\mathbb{E}\Big[\|X^{x+\varepsilon h} - X^{x+\delta h}\|_\infty^{2p}\Big]^{\frac{1}{2}}\cdot\mathbb{E}\Big[\|(K')^\varepsilon\|_\infty^{2p}\Big]^{\frac{1}{2}}.
\end{aligned}$$

Secondly,

$$\begin{aligned}
\text{for (10.2.5)} \Rightarrow & \frac{p}{2}\mathbb{E}\Big[\int_0^T |L_s|^{p-2}\Big|\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_s^\varepsilon\Big) - \nabla_x\sigma\Big(s,\omega,\Xi_s^\delta\Big)d\xi\big](K')_s^\varepsilon\Big|^2 ds\Big],\\
&\le \frac{\mathbb{E}\Big[\|L\|_\infty^p\Big]}{n} + \Big[\frac{n(p-2)}{2}\Big]^{\frac{p-2}{2}}(LT)^{\frac{p}{2}}\mathbb{E}\Big[\|X^{x+\varepsilon h} - X^{x+\delta h}\|_\infty^{2p}\Big]^{\frac{1}{2}}\cdot\mathbb{E}\Big[\|(K')^\varepsilon\|_\infty^{2p}\Big]^{\frac{1}{2}}.
\end{aligned}$$

The terms from (10.2.7) are treated in exactly the same way.

Finally, we use that $\mathbb{E}\big[\|(K')^\varepsilon\|_\infty^{2p}\big]^{1/2} < \infty$ and $\mathbb{E}\big[\|X^{x+\varepsilon h} - X^{x+\delta h}\|_\infty^{2p}\big]^{1/2} \lesssim |\delta-\varepsilon|^p|h|^p$ to conclude

$$\mathbb{E}\Big[\|(K')^\varepsilon - (K')^\delta\|_\infty^p\Big] \lesssim |\varepsilon-\delta|^p|h|^p.$$

Hence by Kolmogorov Continuity Criterion, we have the map $\varepsilon \mapsto (K')_t^\varepsilon(\omega)$ is almost surely continuous for any $t\in[0,T]$ $\mathbb{P}$-almost surely.

Now, we return to Equation (10.2.2). Using the almost sure continuity of $\varepsilon \mapsto X_t^{x+\varepsilon h}(\omega)$ and Assumption 10.2.1 (iv), we have that

$$\varepsilon \mapsto \exp\Bigg(\int_0^t \Big[\int_0^1 \nabla_x b\Big(s,\omega,X_s^x + \xi[X_s^{x+\varepsilon h}(\omega) - X_s^x(\omega)]\Big)d\xi\Big]ds\Bigg),$$

is almost surely continuous. Hence $\varepsilon \mapsto K_t^\varepsilon(\omega)$ is also almost surely continuous. □

Theorem 10.2.5. *Let $p \geq 2$. Let X^x be solution to the SDE* (9.1.1) *under Assumption 10.2.1 and with initial condition $x \in \mathbb{R}^d$. Then we have that $\forall t \in [0,T]$*

$$\frac{X_t^{x+\varepsilon h}(\omega) - X_t^x(\omega)}{\varepsilon} \to h \cdot J_t(\omega) \quad \mathbb{P}\text{-almost surely as } \varepsilon \to 0.$$

Proof. Let $t \in [0,T]$. First, we show convergence in probability of $(X_t^{x+\varepsilon h} - X_t^x)/\varepsilon$ to $h \cdot J_t$ using Proposition 9.2.1. Convergence in probability will imply the existence of a subsequence which converges almost sure. Finally, using Proposition 10.2.4 we know the limit will be almost surely unique.

Writing out the SDE for the increments' process, we have

$$\begin{aligned}
&\frac{X_t^{x+\varepsilon h} - X_t^x}{\varepsilon} - hJ_t \\
&= \int_0^t \Big[\int_0^1 \nabla_x b(s,\omega,\Xi_s^\varepsilon) d\xi - \nabla_x b(s,\omega,X_s^x) \Big] \Big[\frac{X_s^{x+\varepsilon h} - X_s^x}{\varepsilon} \Big] ds \qquad (10.2.9)\\
&+ \int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega,\Xi_s^\varepsilon) d\xi - \nabla_x \sigma(s,\omega,X_s^x) \Big] \Big[\frac{X_s^{x+\varepsilon h} - X_s^x}{\varepsilon} \Big] dW_s \qquad (10.2.10)\\
&+ \int_0^t \nabla_x b(s,\omega,X_s^x) \Big[\frac{X_s^{x+\varepsilon h} - X_s^x}{\varepsilon} - hJ_s \Big] ds \\
&+ \int_0^t \nabla_x \sigma(s,\omega,X_s^x) \Big[\frac{X_s^{x+\varepsilon h} - X_s^x}{\varepsilon} - hJ_s \Big] dW_s,
\end{aligned}$$

where $\Xi^\varepsilon = X^x + \xi[X^{x+\varepsilon h} - X^x(\cdot)]$. As with Theorem 9.4.4, we argue that the terms (10.2.9) and (10.2.10) converge in probability to 0, then use Proposition 9.2.1 to conclude that

$$\Big\| \frac{X_\cdot^{x+\varepsilon h}(\omega) - X_\cdot^x(\omega)}{\varepsilon} - hJ_\cdot(\omega) \Big\|_\infty \xrightarrow{\mathbb{P}} 0.$$

Thus there exists a sequence ε_n such that $\varepsilon_n \to 0$ as $n \to \infty$ and an event $C_1 \subset \Omega$ with $\mathbb{P}[C_1] = 0$ such that $\forall \omega \in \Omega \backslash C_1$

$$\lim_{n\to\infty} \Big\| \frac{X^{x+\varepsilon_n h}(\omega) - X^x(\omega)}{\varepsilon_n} - hJ(\omega) \Big\|_\infty \to 0.$$

Finally, by Proposition 10.2.4 there exists an event $C_2 \subset \Omega$ with $\mathbb{P}[C_2] = 0$ such that $\forall \omega \in \Omega \backslash C_2$ the map

$$\varepsilon \mapsto \Big\| \frac{X^{x+\varepsilon h}(\omega) - X^x(\omega)}{\varepsilon} - hJ(\omega) \Big\|_\infty,$$

is continuous for ε at 0. Then for $\forall \omega \in \Omega \backslash (C_1 \cup C_2)$

$$\lim_{\varepsilon\to 0} \Big\| \frac{X^{x+\varepsilon h}(\omega) - X^x(\omega)}{\varepsilon} - hJ(\omega) \Big\|_\infty \to 0,$$

and $\mathbb{P}[C_1 \cup C_2] = 0$.

□

10.2.3 Invertibility of the Jacobian Matrix

Next, we wish to show that the Jacobian Matrix J_t is $\mathbb{P}$-almost surely invertible for any choice of $t \in [0,T]$. Notice that due to the initial condition, we have that this is true for $t = 0$ since $J_0 = I_d$.

To prove the Jacobian is invertible, we consider a matrix valued stochastic process and observe that for any choice of $t \in [0,T]$, this process will take value equal to the left inverse of J. This proof follows that of Nualart, [Nua06, Chapter 2.3; Equation 2.8].

We introduce the SDE

$$K_t = I_d - \int_0^t K_s\Big[\nabla_x b(s,\omega,X_s) - \Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^{d'}}\big(s,\omega,X_s\big)\Big]ds \\ - \int_0^t K_s \nabla_x \sigma(s,\omega,X_s)dW_s. \tag{10.2.11}$$

Proposition 10.2.6. *Let $p \geq 2$. Let X_x be solution to the SDE* (9.1.1) *under Assumption 10.1.2 and with initial condition $x \in \mathbb{R}^d$. Then we have the following identity $K_tJ_t = I_d$ for all $t \in [0,T]$ $\mathbb{P}$-a.s.*

Proof. We deal here with matrix valued processes which cannot necessarily be assumed commutative, this makes the analysis slightly more involved. Itô's formula for matrices gives that $(KJ)_0 = I_d$ and

$$\begin{aligned} d(KJ)_t =&K_t dJ_t + dK_t J_t + d[K,J]_t, \\ =&K_t\nabla_x b(t,\omega,X_t)J_t dt + K_t\sigma(t,\omega,X_t)J_t dW_t \\ &- K_t\nabla_x b(t,\omega,X_t)J_t dt - K_t\sigma(t,\omega,X_t)J_t dW_t \\ &+ K_t\Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^{d'}}\big(s,\omega,X_s\big)J_t dt \\ &- K_t\Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^{d'}}\big(s,\omega,X_s\big)J_t dt = 0dt + 0dW_t. \end{aligned}$$

□

SDE (10.2.11) does not necessarily satisfy Assumption 9.1.4, the issue being that the term $-z^T\nabla_x b(t,\omega,X_t)z$ is not bounded from above by a constant almost surely for any choice of vector $|z| = 1$. However, an explicit solution to the SDE can be written out pathwise, even if it does not have finite moments. This construction has the property that it is the left inverse of J.

Proposition 10.2.7. *The determinant of the Matrix J_t, denoted D_t, is called the Stochastic Wronskian and satisfies the SDE*

$$\begin{aligned} dD_t =&Tr\Big(\nabla_x b(t,\omega,X_t)\Big)D_t dt + Tr\Big(\nabla_x\sigma(t,\omega,X_t)\Big)D_t dW_t \\ &+ \Big[\Big\langle Tr\big(\nabla_x\sigma(t,\omega,X_t)\big), Tr\big(\nabla_x\sigma(t,\omega,X_t)\big)\Big\rangle_{\mathbb{R}^{d'}} \\ &\qquad - Tr\Big(\big\langle \nabla_x\sigma(t,\omega,X_t), \nabla_x\sigma(t,\omega,X_t)\big\rangle_{\mathbb{R}^{d'}}\Big)\Big]D_t dt, \end{aligned} \tag{10.2.12}$$

with $D_0 = 1$. D_t has explicit form

$$D_t = \exp\Bigg(\int_0^t Tr\Big(\nabla_x b(s,\omega,X_s)\Big) - \frac{1}{2}Tr\Big(\Big\langle \nabla_x\sigma(s,\omega,X_s), \nabla_x\sigma(s,\omega,X_s)\Big\rangle_{\mathbb{R}^{d'}}\Big)ds \\ + \int_0^t Tr\Big(\nabla_x\sigma(s,\omega,X_s)\Big)dW_s\Bigg). \tag{10.2.13}$$

Proof. The proof can be found in [Mao08, Theorem 3.2.2]. The proof involves applying Itô's formula to the determinant of J_t and establishing that it satisfies Equation (10.2.12). Then one applies Itô's formula to Equation (10.2.13) and verifies that this likewise satisfies (10.2.12). Finally, by Theorem 9.1.5, the solution is unique. □

The matrix $\nabla_x b$ being lower semi-definite means that $\mathrm{Tr}(\nabla_x b)$ is bounded from above, but not necessarily from below. We can conclude the $D.$ is almost surely positive and therefore the process K is $\mathbb{P}$-almost surely the inverse (left or right) of J provided $\mathrm{Tr}(\nabla_x b) \neq -\infty$ with positive probability.

Chapter 11

Applications of Malliavin and Parametric Differentiability

In this Chapter, we recover and discuss some standard applications of Malliavin Differentiation and evaluate some of the problems that occur under our framework. In particular, we study the conditions required to prove absolue continuity of the law of a stochastic differential equation (with respect to the Lebesgue measure) (see Theorem 11.1.2) and an integration by parts formula (see 11.2.1).

The results of this Chapter can be found published in [IdRS19, Section 5].

11.1 Representation formulae

Firstly, we present a way of writing the Malliavin Derivative of X^θ in terms of the Jacobian.

Proposition 11.1.1 (Representation formulae)**.** *Let X^x be solution to the SDE* (9.1.1) *under Assumption 9.3.1 and with initial condition $X_0^x = x \in \mathbb{R}^d$. Let J satisfy the SDE* (10.2.1)*. Consider the SDE for the process $J_t J_s^{-1}$ for $t > s$.*

$$\begin{aligned}
J_{s,t} =& J_t J_s^{-1}\\
=& J_s J_s^{-1} + \int_s^t \nabla_x b(r,\omega,X_r(\omega)) J_r J_s^{-1} dr + \int_s^t \nabla_x \sigma(r,\omega,X_r(\omega)) J_r J_s^{-1} dW_r\\
=& I_d + \int_s^t \nabla_x b(r,\omega,X_r(\omega)) J_{s,r} dr + \int_s^t \nabla_x \sigma(r,\omega,X_r(\omega)) J_{s,r} dW_r. \qquad (11.1.1)
\end{aligned}$$

Equation (11.1.1) *is the Fundamental Matrix of the Linear Stochastic Differential Equation* (9.3.1)*. As such, under Assumption 9.3.1 the Malliavin Derivative of X can be expressed for $t > s$ as*

$$D_s X_t = J_{s,t} A(s,t),$$

where $A(s,t)$ is defined for $t > s$ as

$$\begin{aligned}
A(s,t) = \sigma(s,\omega,X_s(\omega)) &+ \int_s^t J_{s,r}^{-1} \Big(U(s,r,\omega) - \Big\langle \nabla_x \sigma(r,\omega,X_r(\omega)), V(s,r,\omega) \Big\rangle_{\mathbb{R}^{d'}} \Big) dr\\
&+ \int_s^t J_{s,r}^{-1} V(s,r,\omega) dW_r.
\end{aligned}$$

Proof. The proof of this representation formula follows the same ideas as Theorem 9.4.1. Equation (9.3.1) is an infinite dimensional SDE, so we project from the infinite dimensional space into a finite dimensional space. We follow the method of [Mao08, Theorem 3.3.1] to solve the solution explicitly in the projection space then use the Dominated Convergence Theorem to ensure the passage to the limit. □

11.1.1 Absolute Continuity

In [Nua06, Theorem 2.3.1], it is proved that the solution of a Stochastic Differential Equation with Lipschitz, deterministic coefficients and elliptic diffusion term has a law which is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$. This proof can be easily extended to the case where the drift term has monotone growth.

Theorem 11.1.2. *Let X^x be solution to the SDE* (9.1.1) *under Assumption 9.3.1 and with initial condition $X_0^x = x \in \mathbb{R}^d$. Suppose additionally that $\forall z \in \mathbb{R}^d$ that*

$$z^T A(s,t) A(s,t)^T z > \lambda(s,t)|z|^2 \geq 0, \qquad \int_0^t \lambda(s,t) ds > 0 \quad \mathbb{P}\text{-almost surely.}$$

Then the law of X_t^x is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$.

Proof. For this proof, recall [Nua06, Corollary 2.1.2] and following that our strategy is to show that the Malliavin matrix is $\mathbb{P}$-almost surely non zero.

The Malliavin Matrix, Q_t is defined to be

$$Q_t = \int_0^t D_s X_t D_s X_t^T ds = J_t \left(\int_0^t K_s A(s,t) A(s,t)^T K_s^T ds \right) J_t^T.$$

Therefore, for $z \in \mathbb{R}^d$ we have $z^T Q_t z \geq \int_0^t \lambda(s,t)|K_s|^2 ds \cdot |J_t|^2 \cdot |z|^2$ which is greater than zero because $|J|, |K| > 0$. □

Remark 11.1.3. *Observe that the Ellipticity condition for σ is no longer enough to ensure that the law is absolutely continuous. When b and σ are deterministic, U and V are uniformly 0 and Ellipticity is enough.*

11.2 Bismut-Elworthy-Li formula

In [Elw92], the author uses Malliavin Differentiability of an SDE X^x to prove differentiability for functions of the form $u(x) = \mathbb{E}[\phi(X_t^x)]$ where ϕ is assumed to be a continuous function and $t \in [0,T]$. This was later extended in [FLL+99] and [FLLL01] to cover functions ϕ which are integrable and even measurable (provided u remains finite).

Define for $t \in (0,T]$ the set $\Gamma_t = \left\{ a \in L^2([0,T]); \int_0^t a_s ds = 1 \right\}$.

Theorem 11.2.1 (Bismut-Elworthy-Li formula)**.** *Let $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ be a bounded, measurable function. Let X^x be solution to the SDE* (9.1.1) *under Assumption 9.3.1 and with initial condition $X_0^x = x \in \mathbb{R}^d$. Let $t \in (0,T]$. Suppose additionally that ($\delta(\cdot)$ stands for the usual Skorokhod integral, see* [Nua06]*)*

1. *$\forall s \in [0,t]$ the matrix $A(s,t)$ has a right inverse,*

2. *$\exists a \in \Gamma_t$ such that $a.A(\cdot,t)^{-1}J. \in dom(\delta)$.*

Then

$$\nabla_x \mathbb{E}\Big[\Phi(X_t^x)\Big] = \mathbb{E}\Big[\Phi(X_t^x)\delta\Big(a.A(\cdot,t)^{-1}J.\Big)\Big].$$

Proof. We give only a sketch of the proof. For a more detailed proof, see [FLL+99] and [FLLL01]. First suppose that Φ is continuously differentiable with bounded derivatives, then

$$\nabla_x \mathbb{E}\Big[\Phi(X_t^x)\Big] = \mathbb{E}\Big[\nabla_x \Phi(X_t^x)\Big] = \mathbb{E}\Big[\nabla \Phi(X_t^x) J_t\Big] = \mathbb{E}\Big[\nabla \Phi(X_t^x) D_s X_t^x A(s,t)^{-1} J_s\Big].$$

Multiplying both sides by $a \in \Gamma_t$, integrating over $[0,t]$ (using $\int_0^t a_s ds = 1$ on the LHS) and

Fubini gives

$$\begin{aligned}\nabla_x \mathbb{E}\Big[\Phi(X_t^x)\Big] &= \mathbb{E}\Big[\int_0^t a_s \nabla\Phi(X_t^x) D_s X_t^x A(s,t)^{-1} J_s ds\Big] \\ =& \mathbb{E}\Big[\int_0^t D_s\Big(\Phi(X_t^x)\Big) a_s A(s,t)^{-1} J_s ds\Big] = \mathbb{E}\Big[\Phi(X_t^x)\delta\Big(a_{.} A(\cdot,t)^{-1} J_{.}\Big)\Big],\end{aligned}$$

where in the last line we used integration-by-parts formula.

Secondly, let Φ be bounded and measurable. Then using that C_b^1 is dense in the set of bounded measurable functions, we approximate Φ by a sequence of functions $\Phi_n \in C_b^1$. Finally, using a domination argument it is shown that one can swap the limits and integrals and one reaches the conclusion. $\square$

Part IV

The Support of McKean-Vlasov Equations

Chapter 12

Functional Quantization

Before tackling the methods to represent the support of McKean-Vlasov equations we address, separately and of independent interest, the *Quantization problem for the law of a Brownian motion as a measure over the collection of Hölder continuous rough paths*. The quantization problem for Gaussian measures for Hilbert spaces was first studied in [LP02], but for Banach spaces, the problem is more challenging with the optimal rate of convergence solved in [GLP03] and separately [DFMS03]. These methods rely on the small ball probabilities of Brownian motion, see [BR92], a tool to measure the compactness of the reproducing kernel Hilbert space unit ball contained in the Banach space.

Using functional analytic methods, we construct a quantization for the law of the Brownian motion that has a rate of convergence that is asymptotically equivalent to the optimal rate of convergence. Our quantization is not optimal, indeed such a quantization does not exist due to the non-weak compactness of the Hölder unit ball. We choose to sacrifice optimality in order to retain certain key properties that allow us to estimate the law of the quantization accurately. To do this, we construct a Karhunen Loève expansion that optimally approximates the Brownian motion with respect to the Hölder norm. Although this representation for Brownian motion is well documented [HIPP14], it is not so well known that the wavelet representation comes from the spectral decomposition of the covariance kernel and so it embodies the optimal approximation by a finite dimensional Gaussian. These quantizations are then enhanced to rough paths and we prove that the rate of convergence for the enhanced quantization to the enhanced Brownian motion is asymptotically the same.

Quantization for rough paths has been first studied in [PS11]. The choice of Karhunen Loève expansion and method of construction used in this work, namely the trigonometric functions, best suits approximations of Brownian motion in the $L^2([0,T];\mathbb{R}^{d'})$ norm. Although this is enough to ensure convergence in the Hölder norm, it is far from efficient and (to the best of our knowledge) no literature exists for rates of convergence. Our approach is demonstrated to be arbitrarily close to optimal and we provide upper and lower bounds on the rate of convergence.

The goal of this Chapter is to construct a finite support measure that approximates the law of an enhanced Brownian motion as a measure over the space of geometric rough paths. The results of this chapter can be found in the preprint [CRS19, Section 3].

12.1 Truncation of Brownian Motion

Using the Cielsielski representation for Brownian motion from Equation (A.1.3), we can obtain a finite dimensional Gaussian measure on $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$

$$W_t^N = \sum_{(p,m)\in\Lambda_N} W_{pm} G_{pm}(t), \tag{12.1.1}$$

which approximates Brownian motion. Let us briefly describe some of the properties this random variable:

- W^N is a Gaussian measure on $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ with Reproducing Kernel Hilbert space

$\mathcal{H}_N = \left(span_{(p,m)\in\Lambda_N}\{G_{pm}\}\right)^{\times d'}$ where G_{pm} are the Schauder functions defined in Definition A.1.1.

- As a finite dimensional Gaussian, the support of W^N is just $\mathcal{H}_N$. This is equal to the space of d'-dimensional, piecewise linear paths over the dyadic intervals of size $T2^{-N}$.
- This is the optimal finite dimensional approximation of Brownian motion with respect to the Hölder norm (A.1.2).
- The support of the measure $\mathcal{L}^{W^N}$ is Reflexive, so by Theorem B.1.2 a stationary quantization exists.

12.1.1 Optimality of the Truncation

We prove that the truncation chosen in (12.1.1) is the optimal choice with respect to the α-Hölder norm. This is an application of the results of [BC19].

Proposition 12.1.1. *Let $\mathcal{L}^W$ be the law of Brownian motion over the Banach space $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$. Then the $d'\cdot 2^N$ dimensional projection $P: C^{\alpha,0}([0,T];\mathbb{R}^{d'}) \to C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ that minimises the integral*

$$\mathbb{E}\left[\left\|W - P[W]\right\|_{\alpha-H\ddot{o}l}^2\right],$$

is the projection

$$P[W]_t = \sum_{(p,m)\in\Lambda_N} W_{pm} G_{pm}(t).$$

Proof. Define the Covariance Kernel $\mathcal{S}: C^{\alpha,0}([0,T];\mathbb{R}^{d'})^* \to C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ by

$$\mathcal{S}[f]_t = \mathbb{E}\Big[f(W)\cdot W_t\Big], \quad \mathcal{S} = \mathfrak{i}\mathfrak{i}^*$$

where $\mathfrak{i}\mathfrak{i}^*$ is the Spectral representation of $\mathcal{S}$. For a Hilbert space $\mathcal{H}$ and a Banach space E, we define the Operator l-topology, on the collection of bounded linear operators $i: \mathcal{H} \to E$ to be

$$l(\mathfrak{i}) := \mathbb{E}\left[\Big\|\sum_{k\in\mathbb{N}} \mathfrak{i}[h_k]\xi_k\Big\|_E^2\right]^{1/2}$$

where $(h_k)_{k\in\mathbb{N}}$ is an orthonormal basis of $\mathcal{H}$ and ξ_k are i.i.d normal random variables. It is well known, see for example [FTJ79], that the closure in the l-topology of the finite rank operators is the compact operators. We wish to find the finite dimensional operator that best approximates the Spectral representation $\mathfrak{i}$ of the Covariance Kernel $\mathcal{S}$ of Brownian motion in the l-topology.

We follow the methods of [BC19]. Using Theorem A.1.2, we can equivalently think of $\mathcal{L}^W$ as a law over the Banach spaces of sequences $(W_{pm})_{(p,m)\in\Delta}$ that satisfy

$$\sup_{(p,m)\in\Delta} 2^{p(\alpha-1/2)}|W_{pm}| < \infty, \quad \lim_{p\to\infty} 2^{p(\alpha-1/2)} \sup_{m=1,\dots,2^p} |W_{pm}| = 0.$$

Equivalently, we think of elements of the dual space $C^{\alpha,0}([0,T];\mathbb{R}^{d'})^*$ as being sequences over Δ that satisfy

$$f = (f_{pm})_{(p,m)\in\Delta}, \quad \|f\|_{\alpha-\text{Höl},*} = \sum_{(p,m)\in\Delta} 2^{p(1/2-\alpha)}|f_{pm}| < \infty.$$

It is well known that we work with the operator

$$\mathfrak{i}(f) = \sum_{(p,m)\in\Delta} f_{pm} W_{pm},$$

where $W_{pm} = \int_0^T W_{pm}(s) H_{pm}(s) ds$ are independent normally distributed random variables with mean 0 and variance 1.

Thus

$$\mathcal{S}\Big[(f_{pm})_{(p,m)\in\Delta}\Big](t) = \sum_{(p,m)\in\Delta} f_{pm}G_{pm}(t), \quad f\Big(\mathcal{S}[f]\Big) = \sum_{(p,m)\in\Delta} |f_{pm}|^2.$$

We wish to maximise this Quadratic form subject to the linear condition

$$\|f\| = \sum_{(p,m)\in\Lambda} 2^{p(1/2-\alpha)}|f_{pm}| = 1.$$

By a simple convexity argument, the functionals that attain this maximisation problem will be wavelet evaluation functionals and hence

$$\lambda^{(001)} = \sup_{\|f\|_{\alpha-\text{Höl},*}=1} f\Big(\mathcal{S}[f]\Big) = 1 = f^{(001)}\Big(\mathcal{S}[f^{(001)}]\Big),$$

where $f^{(001)} = (f^{(001)}_{pm})_{(p,m)\in\Delta}$ satisfies $f^{(001)}_{00} = e_1$ and $f^{(001)}_{pm} = 0$ else. We label $\mathcal{S}[f^{(001)}](t) = x^{(001)}(t) = G_{00}(t)e_1 \in C^{\alpha,0}([0,T];\mathbb{R}^{d'})$. We define $\mathcal{S}_{\mu_{001}}[f] := \lambda^{(001)}f(x^{(001)})x^{(001)}$ and $\mathcal{S}_{001}[f] = \mathcal{S}[f] - \mathcal{S}_{\mu_{001}}[f]$.

By construction, we have that the operator $\mathcal{S}_{\mu_{001}}$ is the Covariance Kernel of the 1-dimensional Gaussian measure that best approximates $\mathcal{L}^W$ in mean square. Equivalently, the Spectral representation $\mathcal{S}_{\mu_{001}} = \mathfrak{i}_{\mu_{001}}\mathfrak{i}^*_{\mu_{001}}$ yields that $\mathfrak{i}_{\mu_{001}}$ is the 1-dimensional operator that best approximates $\mathfrak{i}$ in the l-topology.

By repeating this method, we obtain a sequence of so-called "Rayleigh coefficients" and "Rayleigh Functionals" parametrised by $(q,n,i) \in \Delta \times \{1,...,d'\}$ as

$$\lambda^{(qni)} = 2^{q(2\alpha-1)}, \quad f^{(qni)} = (f^{(qni)}_{pm})_{(p,m)\in\Delta}, \quad f^{(qni)}_{pm} = \delta_{p,q}\delta_{m,n}e_i,$$

and elements G_{qn} that are orthonormal in $\mathcal{H}$.

For fixed $N \in \mathbb{N}$, we obtain the first $d' \cdot 2^N$ elements of these sequences. We construct the projection operator $P_N : C^{\alpha,0}([0,T],\mathbb{R}^{d'}) \to C^{\alpha,0}([0,T],\mathbb{R}^{d'})$ defined by

$$P_N[x](t) = \sum_{(q,n)\in\Delta_N} \sum_{i=1}^{d'} f^{(qni)}(x)G_{qn}(t)e_i.$$

Next, we decompose the law $\mathcal{L}^W = \mu_N * \mathcal{L}^W_N$ where $\mu_N = \mathcal{L}^W \circ P_N^{-1}$ and $\mathcal{L}^W_N = \mathcal{L}^W \circ (I - P_N)^{-1}$. μ_N is a 2^N-dimensional multivariate Gaussian distribution. $\mathcal{L}^W_N$ is a Gaussian measure over $C^{\alpha,0}([0,T],\mathbb{R}^{d'})$ with Kernel S_N that satisfies

$$\sup_{f\in C^{\alpha,0}([0,T],\mathbb{R}^{d'})^*} f\Big(\mathcal{S}_N[f]\Big) \leq 2^{(N+1)(\alpha-1/2)}.$$

In particular, for a random variable W with law $\mathcal{L}^W$ we have that random variable

$$P_N[W](t) = \sum_{(p,m)\in\Delta_N} W_{pm}G_{pm}(t)$$

has law μ_N and

$$\sup_{f\in C^{\alpha,0}([0,T],\mathbb{R}^{d'})^*} \mathbb{E}\Big[f\Big(W - P_N[W]\Big)^2\Big] = 2^{(N+1)(2\alpha-1)}.$$

□

12.1.2 Rate of Convergence of the Truncation

We measure the rate of convergence for a truncated Brownian motion with respect to the α-Hölder norm. We point out that the Banach space $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ is not K-convex (see

[Pis89, Definition 2.3]) so consequently the upper and lower bounds of the rate of convergence cannot be the same.

Proposition 12.1.2. *Let W be a Brownian motion as expressed in* (A.1.3) *and let W^N be truncated Brownian motion* (12.1.1)*. Then for $r>1$ we have*

$$c\cdot d'\cdot N^{1/2-\alpha}\cdot 2^{(\alpha-1/2)N}\leq \mathbb{E}\Big[\|W-W^N\|_\alpha^r\Big]^{1/r}\leq C\cdot d'\cdot\sqrt{N}\cdot 2^{(\alpha-1/2)N}, \tag{12.1.2}$$

where the constants c and C dependent only on α and r.

Proof. Using Theorem A.1.2 and the methods of [BR92], we have

$$\mathfrak{B}_W(\varepsilon)=-\log\Big(\mathbb{P}\Big[\|W\|_\alpha<\varepsilon\Big]\Big)\approx d'\cdot\left(\tfrac{\varepsilon}{d'}\right)^{\frac{-1}{1/2-\alpha}}\quad\text{and}\quad\mathfrak{B}_W(\varepsilon)\lesssim\mathfrak{B}_W(2\varepsilon).$$

Then by [LL99, Proposition 4.1], this implies

$$d'\cdot N^{1/2-\alpha}\cdot 2^{(\alpha-1/2)N}\lesssim\mathbb{E}\Big[\sup_{(p,m)\in\Lambda\setminus\Lambda_N}|W_{pm}|^2 2^{p(2\alpha-1)}\Big]^{1/2}\lesssim d'\cdot\sqrt{N}\cdot 2^{(\alpha-1/2)N},$$

as $N\to\infty$ since W^N is a $d\cdot 2^{N+1}$-dimensional Gaussian random variable.

The Gaussian random variables $W-W^N$ can be dominated by W. By using a concentration inequality and a standard hypercontractivity argument, we can find a constant $C=C(r)$ such that

$$\mathbb{E}\Big[\|W-W^N\|_\alpha^r\Big]\leq C(r)\mathbb{E}\Big[\|W-W^N\|_\alpha^2\Big]^{\frac{r}{2}}.$$

Thus the rate of convergence in mean square is equivalent to the rate of convergence for any choice of r.

□

12.1.3 Enhanced Truncated Brownian Motion

Finally, we prove that the rate of convergence of the *enhanced truncated Brownian motion* to the *enhanced Brownian motion* is the same when the process is lifted to a rough path and studied with respect to the inhomogeneous metric.

The rate of convergence for an enhanced piecewise linear approximation of a Brownian motion has already been studied in [FR11]. Our contribution is a sharper rate of convergence.

Proposition 12.1.3. *Let $N\in\mathbb{N}$ and let $M\geq 2$. Let $\mathcal{L}^{W^N}$ be the law of the truncated Brownian motion over the Banach space $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$. Then $\mathcal{L}^{W^N}$ satisfies Assumption C.2.12 hence W^N can be lifted to an enhanced Gaussian rough path $\mathbf{W}^N=S_M(W^N)$ taking values on the Group $G^M(\mathbb{R}^{d'})$ for $M\geq 2$. Further, for the enhanced Brownian motion $\mathbf{W}$ taking values in $G\Omega_\alpha(\mathbb{R}^{d'})$, there exists a constant $C=C(M,d',\alpha)$ such that*

$$\mathbb{E}\Big[\rho_i(\mathbf{W}^N_{s,t},\mathbf{W}_{s,t})^2\Big]\leq CN2^{(2\alpha-1)N}|t-s|^i, \tag{12.1.3}$$

where $i\in\{1,...,M\}$ and ρ_i is the tensor pseudo-metric (C.2.4) *over $T^M(\mathbb{R}^{d'})$.*

Proof. The case $i=1$ is immediate. We address $i=2$ briefly. For $j,k \in \{1,...,d'\}$ and $j \neq k$

$$\begin{aligned}
\mathbb{E}\Big[\Big|\int_s^t \langle W_{s,r}, e_j\rangle \circ d\langle W_r, e_k\rangle - \int_s^t \langle W^N_{s,r}, e_j\rangle \circ d\langle W^N_r, e_k\rangle\Big|^2\Big]&\\
\leq \int_s^t \int_s^t \mathcal{R}_{\langle W-W^N, e_j\rangle}\begin{pmatrix} s, & s\\ u, & v\end{pmatrix} d\mathcal{R}_{\langle W, e_k\rangle}(u,v)&\\
+\int_s^t \int_s^t \mathcal{R}_{\langle W^N, e_j\rangle}\begin{pmatrix} s, & s\\ u, & v\end{pmatrix} d\mathcal{R}_{\langle W^N-W, e_k\rangle}(u,v)&\\
\leq C|t-s|^2 \cdot \mathbb{E}\Big[\|W-W^N\|_\alpha^2\Big]\cdot \mathbb{E}\Big[\|W\|_\alpha^2\Big].&
\end{aligned}$$

Compiling these terms by summing over j and k completes the $i=2$ case.

For $i>2$, we argue by induction. For a word A such that $|A|=i$ and letter $a \in \mathcal{A}$, we have

$$\begin{aligned}
\mathbb{E}\Big[\Big|\int_s^t \langle \mathbf{W}_{s,r}, e_{(A,a)}\rangle \circ d\langle \mathbf{W}_r, e_{(A,a)}\rangle - \int_s^t \langle \mathbf{W}^N_{s,r}, e_{(A,a)}\rangle \circ d\langle \mathbf{W}^N_r, e_{(A,a)}\rangle\Big|^2\Big]&\\
\leq \int_s^t \int_s^t \mathbb{E}\Big[\rho_i(\mathbf{W}_{s,u}, \mathbf{W}^N_{s,u})\rho_i(\mathbf{W}_{s,v}, \mathbf{W}^N_{s,v})\Big] d\mathcal{R}_{\langle W, e_a\rangle}(u,v)&\\
+\int_s^t \int_s^t \mathbb{E}\Big[\langle \mathbf{W}_{s,u}, e_A\rangle \cdot \langle \mathbf{W}_{s,v}, e_A\rangle\Big] d\mathcal{R}_{\langle W^N-W, e_a\rangle}(u,v)&\\
\leq C|t-s|^{i+1}\cdot \mathbb{E}\Big[\|W-W^N\|_\alpha^2\Big],&
\end{aligned}$$

which implies the inductive hypothesis. □

Theorem 12.1.4. *Let $N \in \mathbb{N}$ and $M \geq 2$. Let $r > 1$. Let $\mathbf{W}^N$ be the enhanced truncated Brownian motion and let $\mathbf{W}$ be the enhanced Brownian motion over $G^M(\mathbb{R}^{d'})$. Then*

$$\mathbb{E}\Big[\rho_{\alpha-\textit{Höl}}\big(\mathbf{W}, \mathbf{W}^N\big)^r\Big]^{1/r} \lesssim \sqrt{N}\cdot 2^{(\alpha-1/2)N} \tag{12.1.4}$$

as $N \to \infty$. Also

$$\mathbb{E}\Big[d_{\alpha-\textit{Höl}}(\mathbf{W}, \mathbf{W}^N)^r\Big]^{1/r} \lesssim \max\Big\{\sqrt{N}2^{(\alpha-1/2)N}, \big(\sqrt{N}2^{(\alpha-1/2)N}\big)^{1/M}\Big\}. \tag{12.1.5}$$

Proof. Firstly, it should be clear that we have

$$\mathbb{E}\Big[\rho_{\alpha-\text{Höl}}(\mathbf{W}, \mathbf{1})^2\Big] < C \quad \text{and} \quad \mathbb{E}\Big[\rho_{\alpha-\text{Höl}}(\mathbf{W}^N, \mathbf{1})^2\Big] < C.$$

Then, we apply [FV10b][Theorem 15.24] with Proposition 12.1.3 to get Equation (12.1.4) in the case $r=2$.

For (12.1.5), we use the well known fact that the the identity operator is $\frac{1}{M}$-Hölder continuous from the space of rough paths paired with the Inhomogeneous metric to the space of rough paths paired with the homogeneous metric and $r=2$.

Now for the case $r \neq 2$. Following [Rie17, Corollary 3.2], we can conclude that the pushforward of $d_\alpha(\mathbf{W}, \mathbf{W}^N)$ with respect to the measure $\mathcal{L}^W$ has a Gaussian tail uniformly on N since the covariance of $W - W^N$ can be dominated by the covariance of W. Then we use a hypercontractivity argument to conclude that

$$\mathbb{E}\Big[d_\alpha(\mathbf{W}, \mathbf{W}^N)^r\Big] \leq C(r)\mathbb{E}\Big[d_\alpha(\mathbf{W}, \mathbf{W}^N)^2\Big]^{\frac{r}{2}}.$$

Thus the rate of convergence in mean square is equivalent to the rate of convergence for any choice of r. □

12.2 Quantization of Brownian Motion

We perform a truncation to obtain a finite dimensional Gaussian that represents an optimal finite dimensional approximation of the Brownian motion. Here, we study how the choice of truncation affects the asymptotic rate of convergence of the quantization error.

Remark 12.2.1. *$\mathcal{L}^{W^N}$ is a non-degenerate measure over the (finite dimensional) vector space $(\mathcal{H}_N, \|\cdot\|_\alpha)$. Therefore by Theorem B.1.2 we know that there exists a codebook $\mathfrak{C}_n = \{\mathfrak{c}_1, ..., \mathfrak{c}_n\}$ and a partition $\hat{\mathfrak{S}}_n = \{\hat{\mathfrak{s}}_1, ..., \hat{\mathfrak{s}}_n\}$ of $\mathcal{H}_N$ such that the quantization $\hat{q}_n$ satisfies*

$$\mathbb{W}^{(2)}_{\mathcal{H}_N, \|\cdot\|_\alpha}\left(\mathcal{L}^{W^N}\Big|_{\mathcal{H}_N}, \mathcal{L}^{W^N} \circ \hat{q}^{-1}\Big|_{\mathcal{H}_N} \right) = \mathbb{E}\Big[\|W^N - \hat{q}_n(W^N)\|_\alpha^2 \Big]^{1/2} = \mathfrak{E}_n$$

However, the measure $\mathcal{L}^{W^N}$ is degenerate over the whole space $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ so constructing an optimal quantization becomes analytically problematic.

Definition 12.2.2. *Let $N \in \mathbb{N}$ be fixed for the moment and let $n \in \mathbb{N}$. Let $\mathcal{L}^{\langle W, e_1 \rangle}$ be the law of Brownian motion over $C^{\alpha,0}([0,T];\mathbb{R})$ and let $\mathcal{L}^{\langle W^N, e_1 \rangle}$ be the law of the 1-dimensional truncated Brownian motion with RKHS $\mathcal{H}^{N,(1)}$. Let $\mathfrak{C}_n^{(1)} = \{\mathfrak{c}_1^{(1)}, ..., \mathfrak{c}_n^{(1)}\}$ and $\hat{\mathfrak{S}}_n^{(1)} = \{\hat{\mathfrak{s}}_1^{(1)}, ..., \hat{\mathfrak{s}}_n^{(1)}\}$ be the codebook and partition of the stationary quantization of $\mathcal{L}^{\langle W^N, e_1 \rangle}$ over $\mathcal{H}^{N,(1)}$.*

Let $\mathfrak{C}_n := \big(\mathfrak{C}_n^{(1)}\big)^{\times d'}$ and $\hat{\mathfrak{S}}_n := \big(\hat{\mathfrak{S}}_n^{(1)}\big)^{\times d'}$. Thus $\mathfrak{C}_n$ and $\hat{\mathfrak{S}}_n$ form a quantization of the truncated Brownian motion over $\mathcal{H}^N$ with independent components. Let $P_N : \mathcal{H} \to \mathcal{H}_N$ be the orthogonal projection and let us continuously extend P_N to $\overline{\mathcal{H}}^\alpha$. We define the new partition of $\overline{\mathcal{H}}^\alpha$ to be

$$\mathfrak{s}_i := (P_N)^{-1}\big[\hat{\mathfrak{s}}_i\big], \quad \mathfrak{S}_n := \Big\{ \mathfrak{s}_1, ..., \mathfrak{s}_n \Big\}. \tag{12.2.1}$$

Pairing the partition $\mathfrak{S}_n$ with the codebook $\mathfrak{C}_n$, we obtain a quantization for the truncated Brownian motion over $\overline{\mathcal{H}}^{\alpha-Höl}$.

It is worth noting that the codebook $|\mathfrak{C}_n| = n^{d'}$. We should also emphasise that the quantization constructed in Definition 12.2.2 is not an optimal quantization of the measures $\mathcal{L}^W$ or $\mathcal{L}^{W^N}$ over the whole space. The reason for this approach is that this quantization exists and is solvable.

Lemma 12.2.3. *Let $n, N \in \mathbb{N}$. Let $\mathcal{L}^W$ be the law of a Brownian motion over $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ with quantization q_n as defined in Definition 12.2.2.*

Let $i \neq j \in \{1, ..., d'\}$. Then $\langle q_n(W), e_i \rangle$ and $\langle q_n(W), e_j \rangle$ are independent.

Proof. For any two sets $C, D \in C^{\alpha,0}([0,T];\mathbb{R})$, we have

$$\begin{aligned}
\mathbb{P}\Big[\langle q_n(W), e_i \rangle \in C, \langle q_n(W), e_j \rangle \in D \Big] &= \mathcal{L}^W \Bigg[\Big(\bigcup_{\substack{k \\ \langle \mathfrak{c}_k, e_i \rangle \in C}} \mathfrak{s}_k \Big) \cap \Big(\bigcup_{\substack{l \\ \langle \mathfrak{c}_l, e_j \rangle \in D}} \mathfrak{s}_l \Big) \Bigg] \\
&= \mathcal{L}^{W^N} \Bigg[\Big(\bigcup_{\substack{k \\ \langle \mathfrak{c}_k, e_i \rangle \in C \cap \mathcal{H}^N}} \hat{\mathfrak{s}}_k \Big) \cap \Big(\bigcup_{\substack{l \\ \langle \mathfrak{c}_l, e_j \rangle \in D \cap \mathcal{H}^N}} \hat{\mathfrak{s}}_l \Big) \Bigg] \\
&= \mathcal{L}^{\langle W^N, e_1 \rangle} \Bigg[\bigcup_{\substack{k \\ \mathfrak{c}_k^{(1)} \in C \cap \mathcal{H}^{N,1}}} \hat{\mathfrak{s}}_k^{(1)} \Bigg] \cdot \mathcal{L}^{\langle W^N, e_1 \rangle} \Bigg[\bigcup_{\substack{l \\ \mathfrak{c}_l^{(1)} \in D \cap \mathcal{H}^{N,1}}} \hat{\mathfrak{s}}_l^{(1)} \Bigg] \\
&= \mathbb{P}\Big[\langle q_n(W), e_i \rangle \in C \Big] \cdot \mathbb{P}\Big[\langle q_n(W), e_j \rangle \in D \Big].
\end{aligned}$$

$\square$

12.2.1 Asymptotic rate of convergence for Quantization

Next, we apply Theorem B.2.1 with Proposition 12.1.2 in order to demonstrate the rate of convergence of the quantization we construct.

Proposition 12.2.4. *Let $\mathcal{L}^W$ be the law of a Brownian on $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ and let $\mathcal{L}^{W^N}$ be the law of the truncated Brownian motion. Choose N to satisfy*

$$N \approx \frac{\mathcal{W}\Big(\ln(2^{2\alpha-1}) \cdot \ln(n)^{2\alpha-1}\Big)}{\ln(2^{2\alpha-1})}, \tag{12.2.2}$$

where $\mathcal{W}$ is the Lambert-W function (see [BL16]*), the inverse function of $y = xe^x$.*

Then $\forall r > 1$, the quantization constructed in Definition 12.2.2 satisfies

$$\bigg(\int_{C^{\alpha,0}([0,T];\mathbb{R}^{d'})} \|x - q_n(x)\|_\alpha^r d\mathcal{L}^W(x) \bigg)^{1/r} \lesssim d' \cdot \Big(\log(n) \Big)^{\alpha-1/2} \tag{12.2.3}$$

as $n \to \infty$.

Proof. It should be clear that the partition as defined in Equation (12.2.1) is not the collection of Voronoi sets generated by the codebook $\mathfrak{C}_n$ over $\overline{\mathcal{H}}^{\alpha-\text{Höl}}$. Thus

$$\mathbb{E}\Big[\big\|W - q_n(W)\big\|_\alpha^r \Big]^{1/r} \geq \bigg(\int_{\overline{\mathcal{H}}^\alpha} \min_{\substack{j=1,\dots,n \\ \mathfrak{c}_j \in \mathfrak{C}}} \|x - \mathfrak{c}_j\|_\alpha^r d\mathcal{L}^W(x) \bigg)^{1/r}.$$

We can further improve this lower bound by minimizing over the all possible codebooks $\mathfrak{C}$ which yields the lower bound

$$d' \Big(\log(n) \Big)^{\alpha-1/2} \lesssim \mathfrak{E}_{n^{d'},r}(\mathcal{L}^W) \leq \mathbb{E}\Big[\big\|W - q_n(W)\big\|_\alpha^r \Big]^{1/r}.$$

For the upper bound, we apply Lemma B.1.6 and Proposition 12.1.2 to get

$$\begin{aligned} \mathbb{E}\Big[\big\|W - q_n(W)\big\|_\alpha^r \Big]^{1/r} \leq& \mathbb{E}\Big[\big\|W^N - q_n(W^N)\big\|_\alpha^r \Big]^{1/r} + \mathbb{E}\Big[\big\|W - W^N\big\|_\alpha^r \Big]^{1/r} \\ \lesssim& \mathfrak{B}_{W^N}^{-1}\Big(\log(n) \Big) + \sqrt{N} \cdot 2^{(\alpha-1/2)N}. \end{aligned}$$

By Theorem B.2.1, we have asymptotic upper and lower bounds on the quantization error for both measures $\mathcal{L}^W$ and $\mathcal{L}^{W^N}$.

Due to the symmetric way in which the truncation and the Hölder norm overlap, we have that

$$\mathbb{P}\Big[\|W^N\|_\alpha \leq \varepsilon \Big] \geq \mathbb{P}\Big[\|W\|_\alpha < \varepsilon \Big],$$

or equivalently

$$-\log\Big(\mathbb{P}\Big[\|W^N\|_\alpha \leq \varepsilon \Big] \Big) = \mathfrak{B}_{W^N}(\varepsilon) \leq \mathfrak{B}_W(\varepsilon) = -\log\Big(\mathbb{P}\Big[\|W\|_\alpha < \varepsilon \Big] \Big).$$

This is true for any choice of truncation level N. Taking the inverse of these bijective, increasing functions gives

$$\mathfrak{B}_{W^N}^{-1}(n) \leq \mathfrak{B}_W^{-1}(n).$$

Thus, for any choice of $N \in \mathbb{N}$,

$$\mathbb{E}\Big[\|W - q_n(W)\|_\alpha^r \Big]^{1/r} \lesssim d' \Big(\log(n) \Big)^{\alpha-1/2} + d'\sqrt{N} \cdot 2^{(\alpha-1/2)N}.$$

Finally, we note that the asymptotic relation of Equation (12.2.2) is equivalent to

$$\sqrt{N} \cdot 2^{(\alpha-1/2)N} \approx \Big(\log(n)\Big)^{\alpha-1/2}$$

which yields the conclusion. □

Remark 12.2.5. *We know by results such as* [DFMS18] *that by sampling a Brownian motion in pathspace, the empirical law will be a good approximation for the law of Brownian motion.*

The difference with this method is that sampling produces a convergence in measure type result. Thus we have proved a deterministic *and not probabilistic result.*

12.2.2 Quantization for a Gaussian Rough Paths

For this Section, we explore lifting our quantized Brownian motion to a rough path. Quantization for rough paths was first studied in [PS11]. In their paper, the authors treat the law of Brownian motion as a measure over the Hilbert space $L^2([0,T];\mathbb{R}^{d'})$. In particular, as a measure over a Hilbert space the authors are able to obtain a stationary quantization, see [LP02]. The Karhunen Loève expansion is obtained using an expansion of trigonometric functions and the authors use well understood pathspace results to establish pointwise convergence of the paths followed by convergence in p-variation. The paths of the quantization codebooks are bounded variation, so they can be lifted to a Signature of a rough path and these converge in the rough path metric to the Brownian rough path. To the best of our knowledge, this is the only work studying quantization in a rough path framework so this chapter is new and of independent interest.

We perform quantization for a Brownian rough path with respect to the pathspace Hölder norm. Due to the nature of the L^2 norm with which the quantization is constructed in [PS11], the approximation with respect to the Hölder norm is far from optimal. By contrast, our approximation is arbitrarily close to optimal. In this Section, we prove that this remains true when the study is carried out with respect to the rough path Hölder norm.

As proved in Lemma B.1.4, the sets $\mathfrak{C} \subset \mathcal{H}$ also have a canonical Young integral signature $\mathbf{c} = S_M(\mathfrak{c})$ for each $\mathfrak{c} \in \mathfrak{C}$.

Definition 12.2.6. *Let $M \geq 2$. Let $\mathcal{L}^W$ be the law of a Brownian motion over $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ and let $\mathcal{L}^{\mathbf{W}}$ be the law of the enhanced Brownian motion over $G\Omega_\alpha(\mathbb{R}^{d'})$. Let q_n be the sequence of quantizations as defined in Definition 12.2.2 for the truncated Brownian motion with N chosen to satisfy Equation* (12.2.2) *and codebooks $\mathfrak{C}_n$ and partitions $\mathfrak{S}_n$.*

Define the sets

$$\mathbf{S} = \Big\{\mathbf{s}_1, ..., \mathbf{s}_n\Big\}, \quad \mathbf{s}_i := \overline{\Big\{\mathbf{h} = S_M(h) : h \in \mathfrak{s}_i \cap \mathcal{H}\Big\}}^{\rho_{\alpha\text{-Höl}}}.$$

These form a partition over the space $G\Omega_\alpha(\mathbb{R}^{d'})$ (up to boundary sets of measure 0). Similarly, define the codebook

$$\mathbf{C} = \{\mathbf{c}_1, ..., \mathbf{c}_n\}, \quad \mathbf{c}_i := S_M[\mathfrak{c}_i].$$

By combining the enhanced codebook with the partition $\mathbf{S}_n$, we obtain the enhanced quantization $\mathbf{q}_n : G\Omega_\alpha(\mathbb{R}^{d'}) \to G\Omega_\alpha(\mathbb{R}^{d'})$

$$\mathbf{q}_n(\mathbf{X}) = \mathbf{c}_i \quad \textit{for } \mathbf{X} \in \mathbf{s}_i, \quad \mathbf{q}_n\Big(G\Omega_\alpha(\mathbb{R}^{d'})\Big) = \mathbf{C}_n. \tag{12.2.4}$$

The next result is an extension of Proposition 12.2.4 to the rough path setting. We follow the same methods as in Section 12.1.3.

Proposition 12.2.7. *Let $M \geq 2$. Fix $N, n \in \mathbb{N}$. Let $\mathcal{L}^{\mathbf{W}}$ be the law of the enhanced Brownian motion. Then there exists a constant $C = C(M, d', \alpha)$ such that*

$$\mathbb{E}\Big[\rho_i\big(\mathbf{W}^N_{s,t}, \mathbf{q}_n(\mathbf{W}^N)_{s,t}\big)^2\Big] \leq C\Big(\log(n)\Big)^{2\alpha-1}|t-s|^i. \tag{12.2.5}$$

Proof. The case $i = 1$ is already proved in Proposition 12.2.4. $i > 2$ can be addressed via an induction argument as in Proposition 12.1.3. Therefore, we only prove the case $i = 2$. Thus for $\mathcal{F}_n$ equal to the σ-algebra generated by the partition of $\mathfrak{S}_n$, we have

$$\begin{aligned}
&\mathbb{E}\left[\Big|\int_s^t \langle W^N_{s,r}, e_i\rangle d\langle W^N_r, e_j\rangle - \int_s^t \langle \mathbb{E}[W^N_{s,r}|\mathcal{F}_n], e_i\rangle d\mathbb{E}[W^N_r|\mathcal{F}_n], e_j\rangle\Big|^2\right] \\
\leq & 2\int_s^t\int_s^t \mathbb{E}\left[\Big\langle W^N_{s,r} - \mathbb{E}[W^N_{s,r}|\mathcal{F}_n], e_i\Big\rangle \cdot \Big\langle W^N_{s,u} - \mathbb{E}[W^N_{s,u}|\mathcal{F}_n], e_i\Big\rangle\right] d\mathbb{E}\left[\langle W^N_r, e_j\rangle \cdot \langle W^N_u, e_j\rangle\right] \\
&+ 2\int_s^t\int_s^t \mathbb{E}\left[\langle W^N_{s,r}, e_i\rangle \cdot \langle W^N_{s,u}, e_i\rangle\right] \cdot d\mathbb{E}\left[\Big\langle W^N_r - \mathbb{E}[W^N_r|\mathcal{F}_n], e_j\Big\rangle \cdot \Big\langle W^N_u - \mathbb{E}[W^N_u|\mathcal{F}_n], e_j\Big\rangle\right], \\
\leq & (t-s)^2\mathbb{E}\left[\Big\|W^N - \mathbb{E}[W^N|\mathcal{F}_n]\Big\|_\alpha^2\right] \cdot \mathbb{E}\Big[\|W^N\|_\alpha^2\Big] \leq C(t-s)^2\Big(\log(n)\Big)^{2\alpha-1},
\end{aligned}$$

using Lemma 12.2.3 and the same Young Estimates as in Proposition 12.1.3.

$$\leq (t-s)^2\mathbb{E}\left[\Big\|W^N - \mathbb{E}[W^N|\mathcal{F}_n]\Big\|_\alpha^2\right] \cdot \mathbb{E}\Big[\|W^N\|_\alpha^2\Big] \leq C(t-s)^2\Big(\log(n)\Big)^{2\alpha-1}.$$

□

Theorem 12.2.8. *Let $r > 1$. Let $\mathcal{L}^W$ be the law of Brownian motion on $G\Omega_\alpha(\mathbb{R}^{d'})$ and let $\mathcal{L}^{\mathbf{W}}$ be the law of the of the enhanced Brownian motion over $G\Omega_\alpha(\mathbb{R}^{d'})$. Let $\mathbf{q}_n$ be the sequence of quantizations constructed in Definition 12.2.6. Then*

$$\left(\int_{G\Omega_\alpha(\mathbb{R}^{d'})} \rho_{\alpha-Höl;[0,T]}\Big(\mathbf{X}, \mathbf{q}_n(\mathbf{X})\Big)^r d\mathcal{L}^{\mathbf{W}}(\mathbf{X})\right)^{1/r} \lesssim \Big(\log(n)\Big)^{\alpha-1/2}. \tag{12.2.6}$$

Proof. The lower bound of Equation (12.2.6) is actually immediate from Equation (12.2.3). The $\rho_{\alpha-\text{Höl}}$ metric can be lower bounded by the projection onto the first level of the Signature so that

$$\mathbb{E}\Big[\rho_{\alpha-\text{Höl};[0,T]}\Big(\mathbf{W}, \mathbf{q}_n(\mathbf{W})\Big)^2\Big] \geq \mathbb{E}\Big[\|W - q_n(W)\|_\alpha^2\Big].$$

Also, by Theorem 12.1.4, we know the rate of convergence for

$$\mathbb{E}\Big[\rho_{\alpha-\text{Höl}}(\mathbf{W}, \mathbf{W}^N)^r\Big]^{1/r} \lesssim \sqrt{N} \cdot 2^{(\alpha-1/2)N} \lesssim \Big(\log(n)\Big)^{\alpha-1/2},$$

where N is the dimension of the linear span of the codebook $\mathfrak{C}_n$ and the choice of Equation (12.2.2) provides the second step. It is clear that

$$\mathbb{E}\Big[\rho_{\alpha-\text{Höl}}(\mathbf{W}, \mathbf{1})^2\Big] < C, \quad \mathbb{E}\Big[\rho_{\alpha-\text{Höl}}(\mathbf{W}^N, \mathbf{1})^2\Big] < C.$$

We can then apply [FV10b][Theorem A.13] with Proposition 12.2.7. We remark that although this method has been used to prove the regularity of enhanced Gaussian rough paths before, there is no part of this method that requires the Gaussian structures, only regularity properties in all moments. Thus

$$\mathbb{E}\Big[\rho_{\alpha-\text{Höl};[0,T]}\Big(\mathbf{W}^N, \mathbf{q}_n(\mathbf{W}^N)\Big)^r\Big]^{1/r} \lesssim \Big(\log(n)\Big)^{\alpha-1/2}.$$

□

Chapter 13

McKean-Vlasov Rough Differential Equations

McKean-Vlasov Equations have also been studied in the context of rough paths, the first study being [CL15]. There, the authors treat the measure dependency as a bounded variation Banach valued operator in the drift term. Thus the measure dependency can be calculated using Banach valued Young integrals and there is no need to exploit the rough path structures beyond what is already necessary to incorporate the noise. The authors prove Existence, Uniqueness and a Propagation of Chaos result for McKean-Vlasov Rough Differential Equations of the form

$$dX_t = \sigma(X_t)d\mathbf{W}_t + b(X_t)d\gamma^\mu_t, \quad \mu = \mathcal{L}^X, \quad X_0 = \xi, \quad t \in [0,T], \tag{13.0.1}$$

where the path $\gamma^\mu_t = \int_0^t \mu_s ds$ represents the measure dependency in the drift term. [CL15] includes an explanation as to why the authors were unable to include a measure dependency in the diffusion terms.

Later, in [BCD18] and more recent preprints [BCD20] and [BCD19], the authors develop the new framework of *Probabilistic Rough Paths*. This insightful development encodes the law of the noise into the rough path, allowing the noise to interact with the measure dependencies and opening up the collection of possible diffusion terms to include adequately regular measure dependencies. Other works that study McKean-Vlasov Equations via rough paths include [DFMS18], [CDFM18] and [CN19].

In this Chapter, we address the approach of [CL15] to solve McKean-Vlasov Rough Differential Equations driven by a Brownian rough path. We choose to present this work in the framework of [CL15] to reduce the complexity and avoid obfuscated algebraic argument.

The results of this chapter can be found in the preprint [CRS19, Section 4].

13.1 Controls and the Accumulated p-Variation

In this first Section, we establish a key condition for the integrability of our quantization. For notational simplicity, we denote $p = \frac{1}{\alpha}$.

Definition 13.1.1. *Let $\beta > 0$ and suppose that $\omega : \Delta_T \to \mathbb{R}^+$ is a control (recall Definition C.2.5). We define the Accumulated β-local ω-variation by*

$$\mathbf{M}_\beta(\omega) := \sup_{\substack{D=(t_i)\\ \omega(t_i,t_{i+1})\le\beta}} \sum_{i:t_i\in D} \omega(t_i, t_{i+1}).$$

The Accumulated β-local controls were first introduced in [CLL13]. We are interested in the specific case where the control is induced by a geometric rough path.

Definition 13.1.2. *Let $\beta > 0$. Let $p > 2$ and let $\mathbf{W} \in G\Omega_\alpha(\mathbb{R}^{d'})$. We define the Accumulated*

β-local p-variation of a geometric rough path to a non-negative function defined by

$$\mathbf{M}_{\beta,p}(\mathbf{W}) := \mathbf{M}_{\beta}(\omega_{\mathbf{W},p}).$$

We define the nondecreasing sequence $(\tau_i(\beta, p, \mathbf{W}))_{i\in\mathbb{N}}$ by

$$\tau_0(\beta) = 0, \quad \tau_{i+1}(\beta) = \inf\{t > \tau_i(\beta); \|\mathbf{W}\|^p_{p-var;[\tau_i(\beta),t]} \geq \beta\} \wedge T. \tag{13.1.1}$$

This is sometimes referred to as the Greedy sequence. *Finally, we define the function $\mathbf{N}_{\beta,p,[0,T]}$: $G\Omega_\alpha(\mathbb{R}^{d'}) \to \mathbb{N} \cup \{\infty\}$ given by*

$$\mathbf{N}_{\beta,p,[0,T]}(\mathbf{W}) := \sup\{n \in \mathbb{N} \cup \{0\} : \tau_n(\beta) < T\}.$$

While stopping time arguments become problematic for McKean-Vlasov Equations due to the presence of the measure dependency, we emphasise that the greedy sequence (13.1.1) is dependent only on the driving noise and not the solution.

It is immediate from the definition that $\mathbf{M}_{\beta,p}(\mathbf{W}) \leq \|\mathbf{W}\|^p_{p-var;[0,T]}$. However, when $\mathbf{W}$ is a Gaussian rough path and $p > 2$, we have $|W_{0,T}|^p \leq \|\mathbf{W}_{0,T}\|^p_{cc} \leq \|\mathbf{W}\|^p_{p-var;[0,T]}$ and $W_{0,T} \sim N(0, T)$ so

$$\mathbb{E}\Big[\exp\Big(\|\mathbf{W}\|^p_{p-var;[0,T]}\Big)\Big] = \infty.$$

Remark 13.1.3. *The Accumulated p-variation is a way of restricting the size of the p-variation in the event that the p-variation becomes large. When the p-variation of a Gaussian is large, by far the most probable event is that there is a single large increment of the process. While the p-variation will increase proportionally to this steep increment, the Accumulated β-local p-variation is restricted to partitions where the increments cannot be larger than β so the one increment does not make a proportional contribution.*

The following Proposition is key to the construction of McKean-Vlasov Rough Differential Equations driven by Gaussian processes.

Proposition 13.1.4. *Let $\mathbf{W}$ be a continuous, centred Gaussian rough path that satisfies Assumption C.2.12. Then $\forall \beta > 0$, the random variable $\mathbf{M}_{\beta,p}(\mathbf{W})$ has well defined Moment Generating Function*

$$[0, \infty) \ni \theta \mapsto \mathbb{E}\big[\exp(\theta \mathbf{M}_{\beta,p}(\mathbf{W})\big] < \infty.$$

Proof. See [CLL13, Theorem 6.3] for tail estimates of the law of the Accumulated p-variation. □

The existence of a moment generating function for the Accumulated p-variation of the driving noise for the McKean-Vlasov Rough Differential Equation is a key Assumption of [CL15], see below. In order to prove propagation of chaos of a sequence of measures, the authors prove that the sequence of empirical measures each has a moment generating function and that the empirical laws converge weakly to the law of the driving noise. We verify the quantization also satisfies this condition:

Lemma 13.1.5 ([FV10a]). *Let $\mathcal{L}^W$ be the law of a Brownian motion over $C^{\alpha,0}([0,T]; \mathbb{R}^{d'})$. Let $h_1, ..., h_n$ be a collection of orthonormal elements of $\mathcal{H}$. Let W^n be a finite Karhunen Loève expansion of W generated by the set $\{h_1, ..., h_n\}$ so that*

$$W^n = \mathbb{E}\big[W\big|\mathcal{F}^n\big],$$

where $\mathcal{F}^n$ is the σ-algebra generated by the functionals $f_j = (i^*)^{-1}[h_j]$ for each $j = 1, ..., n$.

Then the Brownian rough path $\mathbf{W} = S_2(W)$ satisfies the martingale formula

$$\mathbb{E}\Big[\log_{\boxtimes}(\mathbf{W}_{s,t})\Big|\mathcal{F}^n\Big] = \log_{\boxtimes}(\mathbf{W}^n_{s,t}), \tag{13.1.2}$$

where $\mathbf{W}^n_{s,t} = S_2(W^n)_{s,t}$.

The martingale formula yields a very brief proof that the quantized Gaussians are adequately integrable. This first Lemma recasts the well known result mentioned earlier in Equation (B.1.1).

Lemma 13.1.6. *Let $\mathcal{L}^W$ be the law of Brownian motion over $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$. Let $\mathcal{F}$ be a sub-σ algebra of the Borel sigma algebra over $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ that is component-wise conditionally independent. Define $\tilde{W} = \mathbb{E}[W|\mathcal{F}]$. Let $\mathbf{W}$ be the Gaussian rough path of $\mathcal{L}^W$ and $\tilde{\mathbf{W}}$ be the lift of the random variable $\tilde{W}$ to a rough path.*

Then, for a constant $C_1 = C_1(d',p)$ dependent only on d' and p, we have

$$\|\tilde{\mathbf{W}}\|^p_{p-var;[0,T]} \leq C_1 \mathbb{E}\Big[\|\mathbf{W}\|^p_{p-var;[0,T]}\Big|\mathcal{F}\Big].$$

Proof. Firstly, we work with the homogeneous norm (C.1.1) for $G^2(\mathbb{R}^{d'})$ rather than the Carnot Caratheodory norm in order to evaluate the increments explicitly.

By component-wise conditional independence (for the 2nd equality) we have

$$\begin{aligned}
&\|\tilde{\mathbf{W}}\|^p_{p-var;[0,T]}\\
&= \sup_{D=(t_i)} \sum_{i:t_i\in D} \left(\sum_{j=1}^{d'} \Big|\langle \mathbb{E}[W_{t_i,t_{i+1}}|\mathcal{F}], e_j\rangle\Big| + \sum_{\substack{j,k=1\\ j\neq k}}^{d'} \Big| \int_{t_i}^{t_{i+1}} \langle \mathbb{E}[W_{t_i,u}|\mathcal{F}], e_j\rangle d\langle \mathbb{E}[W_u|\mathcal{F}], e_k\rangle\Big|^{1/2}\right)^p\\
&= \sup_{D=(t_i)} \sum_{i:t_i\in D} \left(\sum_{j=1}^{d'} \Big|\langle \mathbb{E}[W_{t_i,t_{i+1}}|\mathcal{F}], e_j\rangle\Big| + \sum_{\substack{j,k=1\\ j\neq k}}^{d'} \Big| \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} \langle W_{t_i,u}, e_j\rangle d\langle W_u, e_k\rangle\Big|\mathcal{F}\Big]\Big|^{1/2}\right)^p\\
&\leq d'^{(2p-2)/p} \sup_{D=(t_i)} \sum_{i:t_i\in D} \mathbb{E}\left[\Big(\sum_{j=1}^{d'} \Big|\langle W_{t_i,t_{i+1}}, e_j\rangle\Big| + \sum_{\substack{j,k=1\\ j\neq k}}^{d'} \Big| \int_{t_i}^{t_{i+1}} \langle W_{t_i,u}, e_j\rangle d\langle W_u, e_k\rangle\Big|^{1/2}\Big)^p \Big|\mathcal{F}\right]\\
&\leq d'^{(2p-2)/p} \mathbb{E}\Big[\|\mathbf{W}\|^p_{p-var;[0,T]}\Big|\mathcal{F}\Big],
\end{aligned}$$

where we use a finite dimensional norm equivalence for the first inequality. There is a further multiplicative constant that appears from translating this result back to the Carnot Caratheodory norm which is dependent only on d'. $\square$

This result does not follow immediately via the same convexity argument used in Equation (B.1.1) because the *Expectation of a Group element may not be a Group element itself.*

Proposition 13.1.7. *Let $n, N \in \mathbb{N}$. Let $\mathcal{L}^W$ be the law of a Brownian motion on $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ and let W^N be the truncated Brownian motion. From Definition 13.1.2, let $\tau_i(\beta)$ be the greedy sequence of the Brownian rough path $\mathbf{W}$, let $\overline{\tau}_i(\overline{\beta})$ be the greedy sequence of the enhanced truncated Brownian motion $\mathbf{W}^N = S_2(W^N)$ and let $\tilde{\tau}_i(\tilde{\beta})$ be the greedy sequence of the enhanced quantization $\mathbf{q}_n(\mathbf{W})$ as introduced in Definition 12.2.6. Let $\overline{\beta} = C_1\beta$ and $\tilde{\beta} = C_1\overline{\beta}$ where C_1 is the constant introduced in Lemma 13.1.6.*

Let $\mathbf{N}_{\beta,p,[0,T]}(\mathbf{W})$, $\overline{\mathbf{N}}_{\overline{\beta},p,[0,T]}(\mathbf{W}^N)$ and $\tilde{\mathbf{N}}_{\tilde{\beta},p,[0,T]}(\mathbf{q}_n(\mathbf{W}))$ be the number of elements of each of the respective greedy sequences over the interval $[0,T]$. Then

$$\tilde{\mathbf{N}}_{\tilde{\beta},p,[0,T]}(\mathbf{q}_n(\mathbf{W})) \leq \overline{\mathbf{N}}_{\overline{\beta},p,[0,T]}(\mathbf{W}^N) \leq \mathbf{N}_{\beta,p,[0,T]}(\mathbf{W}).$$

Proof. This proof relies on the choice of quantization, and we choose $q(W^N)$ to be the optimal quantization of the finite dimensional Gaussian random variable W^N as a measure over the set $\mathcal{H}^N$ with independent spatial components, see Lemma 12.2.3. Let $\tilde{\mathcal{F}}$ be the σ-algebra generated by the partition of the quantization $\tilde{\mathcal{F}} = \sigma(\mathfrak{S})$ and let $\overline{\mathcal{F}}$ be the cylindrical sigma algebra generated by the functionals $(\mathrm{i}^*)^{-1}[\mathcal{H}_N]$. Then we have $q(W^N) = \mathbb{E}[W^N|\tilde{\mathcal{F}}]$ and $W^N = \mathbb{E}[W|\overline{\mathcal{F}}]$.

By Lemma 13.1.6, we therefore have that for any subinterval $[s,t]$

$$\|\mathbf{q}_n(\mathbf{W})\|^p_{p-var;[s,t]} \leq C_1\mathbb{E}\Big[\|\mathbf{W}^N\|^p_{p-var;[s,t]}\Big|\tilde{\mathcal{F}}\Big], \quad \|\mathbf{W}^N\|^p_{p-var;[s,t]} \leq C_1\mathbb{E}\Big[\|\mathbf{W}\|^p_{p-var;[s,t]}\Big|\overline{\mathcal{F}}\Big].$$

In particular, for the intervals $[0, \overline{\tau}_1(\overline{\beta})]$ and $[0, \tau_1(\beta)]$ we have

$$\|\mathbf{q}_n(\mathbf{W})\|^p_{p-var;[0,\overline{\tau}_1(\overline{\beta})]} \leq C_1 \mathbb{E}\Big[\|\mathbf{W}^N\|^p_{p-var;[0,\overline{\tau}_1(\overline{\beta})]}\Big|\tilde{\mathcal{F}}\Big] = C_1\overline{\beta},$$
$$\|\mathbf{W}^N\|^p_{p-var;[0,\tau_1(\beta)]} \leq C_1 \mathbb{E}\Big[\|\mathbf{W}\|^p_{p-var;[0,\tau_1(\beta)]}\Big|\overline{\mathcal{F}}\Big] = C_1\beta.$$

However, by definition we also have $\|\mathbf{q}_n(\mathbf{W})\|^p_{p-var;[0,\tilde{\tau}_1(\tilde{\beta})]} = \tilde{\beta}$ and $\|\mathbf{W}^N\|^p_{p-var;[0,\overline{\tau}_1(\overline{\beta})]} = \overline{\beta}$, so we conclude that $0 < \tilde{\tau}_1(\tilde{\beta}) \leq \overline{\tau}_1(\overline{\beta}) \leq \tau_1(\beta)$.

Next, arguing via induction we suppose that $\tilde{\tau}_k(\tilde{\beta}) \leq \overline{\tau}_k(\overline{\beta}) \leq \tau_k(\beta)$. Then

$$\begin{aligned}
\|\mathbf{W}^N\|^p_{p-var;[\overline{\tau}_k(\overline{\beta}),\tau_{k+1}(\beta)\vee\overline{\tau}_k(\overline{\beta})]} \leq& C_1 \mathbb{E}\Big[\|\mathbf{W}\|^p_{p-var;[\overline{\tau}_k(\overline{\beta}),\tau_{k+1}(\beta)\vee\overline{\tau}_k(\overline{\beta})]}\Big|\overline{\mathcal{F}}\Big]\\
\leq& C_1 \mathbb{E}\Big[\|\mathbf{W}\|^p_{p-var;[\tau_k(\beta),\tau_{k+1}(\beta)]}\Big|\overline{\mathcal{F}}\Big] = C_1\beta,\\
\|\mathbf{q}_n(\mathbf{W})\|^p_{p-var;[\tilde{\tau}_k(\tilde{\beta}),\overline{\tau}_{k+1}(\overline{\beta})\vee\tilde{\tau}_k(\tilde{\beta})]} \leq& C_1 \mathbb{E}\Big[\|\mathbf{W}^N\|^p_{p-var;[\tilde{\tau}_k(\tilde{\beta}),\overline{\tau}_{k+1}(\overline{\beta})\vee\tilde{\tau}_k(\tilde{\beta})]}\Big|\tilde{\mathcal{F}}\Big]\\
\leq& C_1 \mathbb{E}\Big[\|\mathbf{W}^N\|^p_{p-var;[\overline{\tau}_k(\overline{\beta}),\overline{\tau}_{k+1}(\overline{\beta})]}\Big|\tilde{\mathcal{F}}\Big] = C_1\overline{\beta}.
\end{aligned}$$

However, $\|\mathbf{W}^N\|^p_{p-var;[\overline{\tau}_k(\overline{\beta}),\overline{\tau}_{k+1}(\overline{\beta})]} = \overline{\beta}$ and $\|\mathbf{q}_n(\mathbf{W})\|^p_{p-var;[\tilde{\tau}_k(\tilde{\beta}),\tilde{\tau}_{k+1}(\tilde{\beta})]} = \tilde{\beta}$ so we conclude $\tilde{\tau}_{k+1}(\tilde{\beta}) \leq \overline{\tau}_{k+1}(\overline{\beta}) \leq \tau_{k+1}(\beta)$.

Next, suppose that $\mathbf{N}_{\beta,p,[0,T]}(\mathbf{W}) = k$ for some $k \in \mathbb{N}$. Then $T < \tau_{k+1}(\beta) \geq \overline{\tau}_{k+1}(\overline{\beta}) \geq \tilde{\tau}_{k+1}(\tilde{\beta})$. Thus k is an upper bound for $\overline{\mathbf{N}}_{\overline{\beta},p,[0,T]}(\mathbf{W}^N)$ and $\tilde{\mathbf{N}}_{\tilde{\beta},p,[0,T]}(\mathbf{q}_n(\mathbf{W}))$. $\square$

Finally, we establish the uniform integrability of the quantizations.

Proposition 13.1.8. *Let $\mathcal{L}^{\mathbf{W}}$ be the law of an enhanced Brownian motion and let $\mathcal{L}^{\mathbf{W}} \circ \mathbf{q}_n^{-1}$ be the law of the quantized Brownian motion.*

Then the Moment Generating function of the Accumulated p-variation of $\mathbf{q}_n(\mathbf{W})$ is well defined and bounded by the Moment Generating function of the Accumulated p-variation of $\mathbf{W}$.

Proof. From [CLL13, Proposition 4.11], we have

$$\beta\mathbf{N}_{\beta,[0,T]}(\omega) \leq \mathbf{M}_\beta(\omega) \leq \beta\Big(2\mathbf{N}_{\beta,[0,T]}(\omega) + 1\Big),$$

for any control ω so the existence of a Moment Generating Function for $\mathbf{N}$ is equivalent to the existence of a Moment Generating Function for $\mathbf{M}$.

Therefore, by Proposition 13.1.4, we have that $\forall\theta, \beta > 0$ that

$$\mathbb{E}\Big[\exp\Big(\theta\mathbf{N}_{\beta,p,[0,T]}(\mathbf{W})\Big)\Big] < \infty.$$

Applying Proposition 13.1.7, we get that

$$\exp\Big(\theta\tilde{\mathbf{N}}_{(C_1)^2\beta,p,[0,T]}\big(\mathbf{q}_n(\mathbf{W})\big)\Big) \leq \exp\Big(\theta\mathbf{N}_{\beta,p,[0,T]}(\mathbf{W})\Big).$$

We take expectations to conclude. $\square$

13.2 Existence, Uniqueness and the Occupation Measure Path

In this Subsection, we overview some of the key details of [CL15] to establish the link between particle systems and McKean-Vlasov Equations and the existence and uniqueness of the solution law of McKean-Vlasov Equations.

The space of measures μ over the metric space (E, d) is not a Banach space. However, a measure can be thought of as a functional over the space of Lipschitz functions on E.

Definition 13.2.1. *For $\mu \in \mathcal{P}_2(E)$, we define $\gamma^\mu \in \mathrm{Lip}^1_*(E)^*$ to be the linear functional such that for any $f \in \mathrm{Lip}^1_*(E)$,*

$$\gamma^\mu[f] = \int_E f d\mu.$$

Similarly, for a collection of measures $(\mu_t)_{t\in[0,T]}$, we define the Occupation measure path γ^μ_t.

First introduced in [CL15], it is further proved that for the law of an SDE μ_t, the Occupation measure path γ^μ is bounded variation in the Banach norm and so has a canonical Young Signature. The existence and uniqueness of a solution to equation (13.2.1) comes immediately from [FV10b, Chapter 12].

Assumption 13.2.2. *Let $\varsigma > \frac{1}{\alpha}, \gamma > 1$ and $M = \lfloor \frac{1}{\alpha} \rfloor$. Let*

$$\sigma \in \mathrm{Lip}^\varsigma \Big(\mathbb{R}^d, L(\mathbb{R}^{d'}, \mathbb{R}^d)\Big) \quad \textit{and} \quad b \in \mathrm{Lip}^\gamma \Big(\mathbb{R}^d, L(\mathrm{Lip}^1_*(G^M(\mathbb{R}^d))^*, \mathbb{R}^d)\Big).$$

Definition 13.2.3. *Suppose b and σ satisfy Assumption 13.2.2. Let $\mu \in \mathcal{P}_1(G\Omega_\alpha(\mathbb{R}^d)$, $\xi \in \mathbb{R}^d$ and $\mathbf{W} = G\Omega_\alpha(\mathbb{R}^{d'})$.*

Then the operator $\Theta_{b,\sigma} : \mathcal{P}_2\Big(G\Omega_\alpha(\mathbb{R}^d)\Big) \times \mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^{d'}) \to G\Omega_\alpha(\mathbb{R}^d)$ maps $(\mu, \xi, \mathbf{W})$ to the rough path that is the solution of the Rough Differential Equation

$$\begin{aligned} &d\mathbf{X}_t = b(X_t)d\gamma^\mu_t + \sigma(X_t)d\mathbf{W}_t, \quad \mathbf{X}_0 = \xi, \\ &(\mu, \xi, \mathbf{W}) \mapsto \Theta(\mu, \xi, \mathbf{W}) = \mathbf{X}. \end{aligned} \tag{13.2.1}$$

13.2.1 Particle Approximations and Finite Support Laws

Firstly, we address the existence and uniqueness of a solution to the system of interacting particles that the McKean-Vlasov Equation models. Let $\mathfrak{C}$ be a codebook for a quantization of the law of the Brownian motion $\mathcal{L}^W$ as a measure over the Banach space $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ containing n elements $\mathfrak{h}^j$. Each $\mathfrak{h}^j$ is a RKHS path. Associated to each path is a component of the probability vector $\mathfrak{p} = (\mathfrak{p}_j)$ such that $\mathfrak{p}_j = \mathcal{L}^W(\mathfrak{s}_j)$ where $\mathfrak{s}_j \in \mathfrak{S}$ is the element of the partition associated to $\mathfrak{h}^j$.

By the nature of $\mathcal{H}$, we know that each path $\mathfrak{h}^j$ is a 1-variation path. Hence one can construct a canonical lift from $\mathfrak{h}^j$ to a rough path $\mathbf{h}^j$ using Young Integration over the interval $[0,T]$. Thus for $t \in [0,T]$ where M is the largest integer such that $M\alpha < 1$ we have

$$\mathbf{h}^j_t = S_M(\mathfrak{h}^j)_{0,t}.$$

We know that n is a finite integer, so we can denote the single path $\mathfrak{h} := \times_{j=1}^n \mathfrak{h}^j$ which takes values in $\mathbb{R}^{d'\times n}$. This path is still 1-variation with respect to the canonical norm on $\mathbb{R}^{d'\times n}$. Therefore, we can similarly construct

$$\mathbf{h}_t = S_M\Big(\times_{j=1}^n \mathfrak{h}^j\Big)_{0,t}.$$

For clarity, we emphasise that this is a rough path taking values in $T^M\Big(\mathbb{R}^{d'\times n}\Big)$ and it is not the same as $\oplus_{j=1}^n \mathbf{h}^j = \oplus_{j=1}^n S_M(\mathfrak{h}^j)$ which takes values in $T^M(\mathbb{R}^{d'})^{\oplus n}$.

When working on the tensor algebra $T^M(V)$. We refer to the Alphabet $\mathcal{A}$, which in the case $V = \mathbb{R}^{d'}$, is the letters $\{1, ..., d'\}$. However, when working on the tensor algebra $T^M\Big(\mathbb{R}^{d'\times n}\Big)$, we have the Alphabet $\mathcal{A}$ containing all the pairs $\big\{(i,j); i \in \{1,...,d'\}, j \in \{1,...,n\}\big\}$. We will also refer to $\mathcal{A}^j$, the Subalphabet containing all pairs $\big\{(i,j); i \in \{1,...,d'\}\big\}$. Key to the following result is that the Subalphabets $\mathcal{A}^j$ form a partition of the Alphabet $\mathcal{A}$.

Lemma 13.2.4. *Let V be a vector space with finite Alphabet $\mathcal{A}$ and suppose that $\mathcal{A}$ can be partitioned into a finite number of Subalphabets denoted by $\mathcal{A}^j$. Define*

$$I^M(V) := \Big\{h \in T^M(V) : \langle h, e_I\rangle = 0, \forall I \textit{ a word with letters in } \mathcal{A} \textit{ s.t. } \exists j \textit{ where } I \textit{ is a word of } \mathcal{A}^j\Big\}.$$

Then $I^M(V)$ is a closed ideal of the Lie Algebra $P^M(V)$.

Proof. We verify that for $h_1 \in I^M(V)$ and $h_2 \in P^M(V)$ that $[h_1, h_2]_{\boxtimes} \in I^M(V)$.

Let I be a word that has the property that $\exists j$ such that I is also a word of $\mathcal{A}^j$. We denote

$$\Delta e_I = \sum_{I_1 I_2 = I} e_{I_1} \otimes e_{I_2}$$

using "Sweedler" notation and $I_1 I_2$ as being word concatenation. If I is a word with letters in $\mathcal{A}^j$ then any subword of I is also a word with letters in $\mathcal{A}^j$.

Therefore, for $h_1 \in I^M(V)$ and $h_2 \in P^M(V)$

$$\langle h_1 \boxtimes h_2, e_I \rangle = \langle h_1 \otimes h_2, \Delta e_I \rangle = \sum_{I_1 I_2 = I} \langle h_1, e_{I_1} \rangle \cdot \langle h_2, e_{I_2} \rangle = \sum_{I_1 I_2 = I} 0 \cdot \langle h_2, e_{I_2} \rangle = 0.$$

Similarly $\langle h_2 \boxtimes h_1, e_I \rangle = 0$, so naturally

$$\Big\langle [h_1, h_2]_{\boxtimes}, e_I \Big\rangle = 0.$$

□

Given an Ideal of a Lie Algebra, one can obtain a normal subgroup of the associated Lie Group by taking exponentials. Thus define

$$K^M(V) := \exp_{\boxtimes}\Big(I^M(V)\Big), \tag{13.2.2}$$

and consider the quotient group $G^M(V)/K^M(V)$. There is a canonical isomorphism that maps this quotient group to $\oplus_j G^M(V^j)$ where V^j is the vector space with Alphabet $\mathcal{A}^j$.

In order to study the system of interacting particle equations for (13.0.1), we consider the following drift and diffusion terms. Before that, we introduce a notational convenience in order to distinguish between elements of $\mathbb{R}^d$ and $\mathbb{R}^{d\times n}$. Recall that for $i \in \mathcal{A}$, e_i is the unit vector in the vector space with Alphabet $\mathcal{A}$. We denote $Y \in \mathbb{R}^{d\times n}$ and $\langle Y, e_{(\cdot,m)}\rangle \in \mathbb{R}^d$ to be the canonical projection of Y where $m \in \{1, ..., n\}$.

Definition 13.2.5. *Let b and σ satisfy Assumption 13.2.2. Let $\mathfrak{p} = (\mathfrak{p}_k)_{k=1,\dots n} \in \mathfrak{P}$. Let $B : \mathbb{R}^{d\times n} \to \mathbb{R}^{d\times n}$ and $\Sigma : \mathbb{R}^{d\times n} \to L\big(\mathbb{R}^{d'\times n}, \mathbb{R}^{d\times n}\big)$ be defined by*

$$\begin{aligned} B(X) :=& \bigoplus_{m=1}^n \left(b\big(\langle X, e_{(\cdot,m)}\rangle\big)\Big[\sum_{k=1}^n \mathfrak{p}_k \delta_{\langle X, e_{(\cdot,k)}\rangle}\Big]\right), \\ \Sigma(X) :=& \operatorname{Diag}_{m=1,\dots,n}\Big(\sigma\big(\langle X, e_{(\cdot,m)}\rangle\big)\Big). \end{aligned}$$

Let $\tilde{\mathbf{W}}^k_t \in G\Omega_\alpha(\mathbb{R}^{d'})$ for each $k \in \{1, ..., n\}$. Let $\tilde{\mathbf{W}} = \bigoplus_{k=1}^n \tilde{\mathbf{W}}^k$ be the rough path taking values in the quotient group $G^M\big((\mathbb{R}^{d'\times n}\big)/I^M\big((\mathbb{R}^{d'\times n}\big)$. Let $\mathbb{X}_t$ be the controlled rough that solves the Rough Differential Equation

$$dX_t = B(X_t)dt + \Sigma(X_t)d\tilde{\mathbf{W}}_t, \quad X_0 = \xi^{\times n}, \tag{13.2.3}$$

taking values in $\mathbb{R}^{d\times n}$. By the properties of b and σ from Definition 13.2.3, we have that $B \in \operatorname{Lip}^\gamma\Big((\mathbb{R}^{d\times n}\Big)$ and $\Sigma \in \operatorname{Lip}^\varsigma\Big((\mathbb{R}^{d\times n}, L\big(\mathbb{R}^{d'\times n}, \mathbb{R}^{d\times n}\big)\Big)$. Therefore, the existence of a solution to Equation (13.2.3) is standard.

Next we introduce a product on the space of vector fields from U into $T^M(V, U)$ designed to simplify the representation of a controlled rough path.

Definition 13.2.6. *Let V and U be vector spaces. Let $i, j \in \mathbb{N}$. For differentiable Vector fields $F : U \to L(V^{\times i}, U)$ and $G : U \to L(V^{\times j}, U)$, we define the operation $\star$ such that $F \star G : U \to$*

$L(V^{\times(i+j)}, U)$ *by*

$$F \star G(u)\Big[v_1, ..., v_j, v_{j+1}, ..., v_{j+i}\Big] = \left(\lim_{\varepsilon \to 0} \frac{F\big(u + \varepsilon G(u)[v_1, ..., v_j]\big) - F\big(u\big)}{\varepsilon}\right)[v_{j+1}, ..., v_{j+i}]$$
$$= DF(u)\Big(G(u)[v_1, ..., v_j]\Big)[v_{j+1}, ..., v_{j+i}] \tag{13.2.4}$$

It is a natural observation to make that the controlled rough path $\mathbb{X}$ that represents the solution to Equation (13.2.3) is equal to

$$\mathbb{X}_s = \Big(X_s, \Sigma(X_s), \Sigma \star \Sigma(X_s), ..., \Sigma^{\star(M-1)}(X_s)\Big), \quad s \in [0, T]. \tag{13.2.5}$$

Lemma 13.2.7. *Let V, U be vector spaces with alphabets $\mathcal{A}$ and $\hat{\mathcal{A}}$. Suppose that $V = \oplus_{j=1}^n V^j$ and $U = \oplus_{j=1}^n U^j$ so that $\mathcal{A}$ and $\hat{\mathcal{A}}$ can be partitioned into a collection of n subalphabets $\mathcal{A}^j$ and $\hat{\mathcal{A}}^j$ for $j = 1, ..., n$.*

For $k, l \in \mathbb{N}$, let $F : U \to L(V^{\oplus k}, U)$ and $G : U \to L(V^{\oplus l}, U)$, suppose that there exist $f^j : U^j \to L\Big((V^j)^{\oplus k}, U^j\Big)$ and $g^j : U^j \to L\Big((V^j)^{\oplus k}, U^j\Big)$ such that we have the representation

$$F(u) = \mathrm{Diag}_{j=1,...,n}\Big(f^j(P_{U^j}[u])\Big), \quad G(u) = \mathrm{Diag}_{j=1,...,n}\Big(g^j(P_{U^j}[u])\Big).$$

Suppose that F is differentiable. Then $F \star G$ has the representation

$$F \star G(u) = \mathrm{Diag}_{j=1,...,n}\Big(Df(P_{U^j}[u]) \times g(P_{U^j}[u])\Big). \tag{13.2.6}$$

Proof. For fixed $m \in \{1, ..., n\}$, let $u_m \in U^m$ and let I be a word of the subalphabet $\mathcal{A}^m$ such that $I = (I_1, I_2)$ where $|I_1| = k$ and $|I_2| = l$.

Outside of this scenario, all derivatives will be 0 by construction. □

We know that by Theorem C.3.3, the controlled rough path $\mathbb{X}$ can be lifted to a rough path. Our next result, the main result of this Section and similar to one found in [CL15], ensures the choice of lift does not affect the final solution to our equations.

Theorem 13.2.8. *For $j = 1, ..., n$, let $\mathbf{W}^j \in G\Omega_\alpha(\mathbb{R}^{d'})$ and define $\mathbf{W} = \oplus_{j=1}^n \mathbf{W}^j$. Let $\tilde{\mathbf{W}}$ be the extension of $\mathbf{W}$ to $G\Omega_\alpha\Big(\mathbb{R}^{d' \times n}\Big)$. Let B and Σ be as defined in Definition 13.2.5 and let $\mathbb{X}$ be the unique controlled rough path that solves the Rough Differential Equation* (13.2.3).

Let $\mathbf{X} \in G\Omega_\alpha\Big(\mathbb{R}^{d \times n}\Big)$ be the lift of $\mathbb{X}$ as constructed in Equation (C.3.2). *Then $\mathbf{X}$ is dependent on $\mathbf{W}$ but not $\tilde{\mathbf{W}}$.*

Proof. Let $V = \mathbb{R}^{d' \times n}$ and $U = \mathbb{R}^{d \times n}$ with alphabets $\mathcal{A}$ and $\hat{\mathcal{A}}$ both with n subalphabets

$$\mathcal{A}^j = \Big\{(i, j) : i \in \{1, ..., d'\}\Big\}, \quad \hat{\mathcal{A}}^j = \Big\{(i, j) : i \in \{1, ..., d\}\Big\}.$$

where $j \in \{1, ..., n\}$. Thus all the vector spaces V^j are isomorphic to $\mathbb{R}^{d'}$ and U^j are isomorphic to $\mathbb{R}^d$ but each V^j and U^j is distinct and identifyable. As with the normal subgroup constructed in Equation (13.2.2), we know the normal subgroup that generates the cosets for the quotient group is

$$I^M\Big(\mathbb{R}^{d' \times n}\Big) = \Big\{h \in T^M\Big(\mathbb{R}^{d' \times n}\Big) : \langle h, e_I \rangle = 0, \forall I \text{ s.t. } \exists j \in \{1, ..., n\} \text{ with } I \in \mathcal{A}^j\Big\},$$
$$K^M\Big(\mathbb{R}^{d' \times n}\Big) = \exp_{\boxtimes}\Big(I^M(\mathbb{R}^{d' \times n})\Big).$$

For $s, t \in [0, T]$, let

$$\pi_{G^M\left(\mathbb{R}^{d' \times n}\right)/K^M\left(\mathbb{R}^{d' \times n}\right)}\Big[\tilde{\mathbf{W}}_{s,t}\Big] = \tilde{\mathbf{W}}_{s,t} \boxtimes K^M\big(\mathbb{R}^{d' \times n}\big) = \mathbf{W}_{s,t}.$$

By Theorem C.3.3, we know this is equal to

$$\mathbf{X}_{s,t} := \mathbf{1} + \lim_{|D|\to 0} \sum_{i:t_i\in D} \left(\sum_{k=1}^{M} \Big\langle \big(\mathbb{X}_{t_i} - X_{t_i}\big)^{\otimes k}, (\delta_{\succ})^k[\mathbf{W}_{t_i,t_{i+1}}] \Big\rangle + B(X_{t_i})(t_{i+1} - t_i) \right), \quad (13.2.7)$$

and $\mathbb{X}$ is defined as in Equation (13.2.5). It is important to realise that the drift term, the only coefficient that is dependent on the law of the solution, is only included in the first level of the signature. Measure dependencies are generally smoother than path dependencies and their higher regularity means they are $o\big(|D|^{1+}\big)$.

Next for $t_i \in D$, we have

$$\begin{aligned}
(\mathbb{X}_{t_i} - X_{t_i})^{\otimes k} =& \Big(\Sigma(X_{t_i}) + \Sigma \star \Sigma(X_{t_i}) + ... + \Sigma^{\star(M-1)}(X_{t_i})\Big)^{\otimes k} \\
=& \sum_{\substack{l_1,...,l_k=1 \\ l_1+...+l_k \le M}}^{M-1} \bigotimes_{m=1}^{k} \Sigma^{\star l_m}(X_{t_i}).
\end{aligned}$$

Using Definition 13.2.5 and Lemma 13.2.7 we have that there exists $f_j : U^j \to L\Big((V^j)^{\oplus l_m}, U^j\Big)$ such that

$$\Sigma^{\star l_m}(X_s) = \mathrm{Diag}_{j=1,...,n}\Big(f_j(\langle X_s, e_{(\cdot,j)}\rangle)\Big).$$

Similarly, there exist functions $g_j : U^j \to L\Big((V^j)^{\oplus(l_1+...+l_k)}, U^j\Big)$ such that

$$\bigotimes_{m=1}^{k} \Sigma^{\star l_m}(X_s) = \mathrm{Diag}_{j=1,...,n}\Big(g_j(\langle X_s, e_{(\cdot,j)}\rangle)\Big),$$

which is an operator restricted to the subgroup $\oplus j = 1^n G^M(V^j)$. Thus Equation (13.2.7) is dependent on the tensor of rough paths $\mathbf{W}$ and not on the Extension $\tilde{\mathbf{W}}$. $\square$

13.2.2 Existence and Uniqueness

For this section, we focus on the approach of [CL15]. Firstly, we introduce some of the notation and operators used in this paper to construct different elements for solving our McKean-Vlasov equation. The methods and results of [BCD18] which are further explored in [BCD20] and [BCD19] are not used here.

Definition 13.2.9. *Let b and σ satisfy Assumption 13.2.2. Let $\mathcal{L} \in \mathcal{P}_1\big(G\Omega_\alpha(\mathbb{R}^{d'})\big)$ and $\mu \in \mathcal{P}_1\big(G\Omega_\alpha(\mathbb{R}^d)\big)$ be probability measures. Then define the map $\Psi_{\mathcal{L}} : \mathcal{P}_2\big(G\Omega_\alpha(\mathbb{R}^d)\big) \to \mathcal{P}_2\big(G\Omega_\alpha(\mathbb{R}^d)\big)$ by*

$$\Psi_{\mathcal{L}}(\mu) = \mathcal{L} \circ \Theta_{b,\sigma}(\mu, \xi, \cdot)^{-1}. \quad (13.2.8)$$

The fixed point of the operator $\Psi_{\mathcal{L}}$ will be the law of the solution to the McKean-Vlasov Equation (13.0.1) where the law of the driving noise $\mathbf{W}$ is given by $\mathcal{L}$.

Assumption 13.2.10. *Let $\varsigma > \frac{1}{\alpha} > 1$ and $\gamma > 1$. Suppose that*

1. *The measure $\mathcal{L}^{\mathbf{W}} \in \mathcal{P}_2(G\Omega_\alpha(\mathbb{R}^{d'}))$ satisfies that for any $\theta \ge 0$*

$$\int_{G\Omega_\alpha(\mathbb{R}^{d'})} \exp\Big(\theta \mathbf{M}_{1,[0,T]}(\omega_{\mathbf{X}})\Big) d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) < \infty, \quad (13.2.9)$$

2. *The functions b and σ satisfy Assumption 13.2.2.*

Theorem 13.2.11 ([CL15]). *Suppose Assumption 13.2.10 holds. Then the operator $\Psi_{\mathcal{L}^{\mathbf{W}}}$ is a contraction operator with fixed point equal to the law of the solution to the McKean-Vlasov Equation* (13.0.1).

Hence there exists a unique solution to the Rough Differential Equation (13.0.1).

13.3 Propagation of Chaos and Quantization

The final result of [CL15] is to prove continuity of the map from the law of the driving noise to the law of the McKean-Vlasov Equation. This is framed within the narrative of “Propagation of Chaos”. We exploit this result to show that the law of the associated particle systems of our quantizations converge to the true law of the McKean-Vlasov Equation.

Definition 13.3.1. *Let $K:(0,\infty)\to(0,\infty)$ be a monotone increasing real valued function. Define the collection of measures*

$$\mathcal{P}_K\Big(G\Omega_\alpha(\mathbb{R}^{d'})\Big) := \Bigg\{\mathcal{L}\in\mathcal{P}_1\Big(G\Omega_\alpha(\mathbb{R}^{d'})\Big) : \forall\theta\in(0,\infty)$$
$$\int_{G\Omega_\alpha(\mathbb{R}^{d'})} \exp\Big(\theta\mathbf{M}_{1,[0,T]}(\omega_{\mathbf{X}})\Big)d\mathcal{L}(\mathbf{X}) \leq K(\theta)\Bigg\}$$

paired with the topology of weak convergence generated by the rough path Hölder norm.

A natural way to think about this collection of measures is the law of all rough paths such that the moment generating function of the Accumulated $\frac{1}{\alpha}$-variation is dominated by the function K.

Proposition 13.3.2. *Suppose Assumption 13.2.10 is satisfied. Suppose additionally that there exists a monotone increasing function $K:(0,\infty)\to(0,\infty)$ that dominates Equation* (13.2.9). *Define the operator $\Xi:\mathcal{P}_K\big(G\Omega_\alpha(\mathbb{R}^{d'})\big)\to\mathcal{P}_2\big(G\Omega_\alpha(\mathbb{R}^{d})\big)$ by*

$$\Xi\big[\mathcal{L}^{\mathbf{W}}\big]=\mathcal{L}^{\mathbf{X}}, \tag{13.3.1}$$

where $\mathcal{L}^{\mathbf{X}}$ is the unique measure that is a fixed point of Equation (13.2.8) *so that $\Psi_{\mathcal{L}^{\mathbf{W}}}(\mathcal{L}^{\mathbf{X}})=\mathcal{L}^{\mathbf{X}}$.*

Then the operator is well defined and for $\mathcal{L}^{\mathbf{W}_1},\mathcal{L}^{\mathbf{W}_2}\in\mathcal{P}_K(G\Omega_\alpha(\mathbb{R}^{d'}))$ we have

$$\mathbb{W}^{(2)}_{\rho_{\alpha-H\ddot{o}l;[0,T]}}\Big(\Xi[\mathcal{L}^{\mathbf{W}_1}],\Xi[\mathcal{L}^{\mathbf{W}_2}]\Big)\leq C\mathbb{W}^{(2)}_{\rho_{\alpha-H\ddot{o}l;[0,T]}}\Big(\mathcal{L}^{\mathbf{W}_1},\mathcal{L}^{\mathbf{W}_2}\Big) \tag{13.3.2}$$

with a constant $C=C(\alpha,K,T,d,d')$.

Previously, this result was used to show that the empirical measure obtained by sampling paths of a Brownian motion could be used to obtain a particle system that would converge as the number of particles increased to the solution of a McKean-Vlasov Equation. In the remarkable work [DFMS18], the authors study the rate of convergence of these empirical measures to the true law in probability.

Proof. Same as proof [CL15, Lemma 4.11]. □

13.4 Continuity with respect to the Occupation Measure path

In [CL15, Theorem 4.9], the goal was to establish the existence of a contraction operator whose fixed point would be the law of the McKean-Vlasov Equation. In fact, computing the specific contraction operator is not simple. Here, we provide a more tangible operator that is (Lipschitz) continuous but not a contraction.

Proposition 13.4.1. *Let b and σ satisfy Assumption 13.2.2 and let $\Theta_{b,\sigma}$ be the operator from Definition 13.2.3.*

Then $\Theta_{b,\sigma}$ is Locally Lipschitz continuous in the measure component, that is $\forall\mu,\nu\in\mathcal{P}_1(G\Omega_\alpha(\mathbb{R}^d))$ such that

$$\int_{G\Omega_\alpha(\mathbb{R}^d)}\rho_{\alpha-H\ddot{o}l}(\mathbf{X},\mathbf{1})d\mu(\mathbf{X}),\int_{G\Omega_\alpha(\mathbb{R}^d)}\rho_{\alpha-H\ddot{o}l}(\mathbf{X},\mathbf{1})d\nu(\mathbf{X})<C$$

and $\forall \xi \in \mathbb{R}^d$ *and* $\forall \mathbf{W} \in G\Omega_\alpha(\mathbb{R}^{d'})$ *such that* $\rho_{\alpha-H\ddot{o}l;[0,T]}(\mathbf{W}, 1) < C$*,* $\exists L_C > 0$ *such that*

$$\rho_{\alpha-H\ddot{o}l}\Big(\Theta_{b,\sigma}(\mu,\xi,\mathbf{W}), \Theta_{b,\sigma}(\nu,\xi,\mathbf{W})\Big) \le L_C \mathbb{W}^{(2)}_{\rho_{\alpha-H\ddot{o}l;[0,T]}}\Big(\mu,\nu\Big) \tag{13.4.1}$$

Proof. Let $p = \frac{1}{\alpha}$ and $M = \lfloor \frac{1}{\alpha} \rfloor$. Denote the control $\omega(s,t) = \|\mathbf{W}\|^p_{p-var;[s,t]} + \|\gamma^\mu\|_{1-var;[s,t]} + \|\gamma^\nu\|_{1-var;[s,t]}$. Then [CL15, Lemma 4.3] gives

$$\rho_{p-\omega;[0,T]}\Big(\Theta_{b,\sigma}(\mu,\xi,\mathbf{W}), \Theta_{b,\sigma}(\nu,\xi,\mathbf{W})\Big) \le C\rho_{1,\omega;[0,T]}(\gamma^\mu,\gamma^\nu)\exp\Big(\mathbf{M}_{\beta,[0,T]}(\omega)\Big).$$

Indeed, we also have

$$\|\gamma^\mu_{s,t} - \gamma^\nu_{s,t}\|_{\mathrm{Lip}^1(G^M(\mathbb{R}^d))^*} \le |t-s|\mathbb{W}^{(2)}_{\rho_{\alpha-\text{Höl};[0,t]}}\Big(\mu,\nu\Big).$$

By assumption, the Wasserstein distance must be finite across the interval $[0,T]$, so we know the control can be dominated by $\omega(s,t) \lesssim |t-s|$. Thus $\rho_{p-\omega}$ will be equivalent to $\rho_{\alpha-\text{Höl}}$ and we get

$$\rho_{\alpha-\text{Höl};[0,T]}\Big(\Theta_{b,\sigma}(\mu,\xi,\mathbf{W}), \Theta_{b,\sigma}(\nu,\xi,\mathbf{W})\Big) \le C\mathbb{W}^{(2)}_{\rho_{\alpha-\text{Höl};[0,t]}}\Big(\mu,\nu\Big)\cdot\exp\Big(\mathbf{M}_{\beta,[0,T]}(\omega)\Big)$$

Next, we note that while the constant C is uniform over the choice of μ and ν, the control ω is dependent on them and so the Accumulated β-local p-variation is also dependent on their second moments. □

With only Proposition 13.4.1, one can establish the distance between two paths driven by different occupation measure paths. Next we prove uniform continuity.

Theorem 13.4.2. *Let* b *and* σ *satisfy Assumption 13.2.2 and let* $\Theta_{b,\sigma}$ *be the operator defined in Definition 13.2.3. Then the operator* $\Theta_{b,\sigma}$ *is jointly continuous over* $\mathcal{P}_2\big(G\Omega_\alpha(\mathbb{R}^d)\big)\times\mathbb{R}^d\times G\Omega_\alpha(\mathbb{R}^{d'})$*. In particular,*

$$\begin{aligned}\lim_{(\mu_k,\xi_k,\mathbf{W}_k)\to(\mu,\xi,\mathbf{W})}\Theta_{b,\sigma}(\mu_k,\xi_k,\mathbf{W}_k) &= \lim_{\mu_k\to\mu}\lim_{\xi_k\to\xi}\lim_{\mathbf{W}_k\to\mathbf{W}}\Theta_{b,\sigma}(\mu_k,\xi_k,\mathbf{W}_k)\\ &= \Theta_{b,\sigma}(\mu,\xi,\mathbf{W})\end{aligned}$$

Proof. Let $\xi,\chi \in \mathbb{R}^d$ and $p = \frac{1}{\alpha}$. For $\mathbf{W}_1, \mathbf{W}_2 \in G\Omega_\alpha(\mathbb{R}^{d'})$ and $\mu,\nu \in \mathcal{P}_2\Big(G\Omega_\alpha(\mathbb{R}^d)\Big)$, define the control

$$\omega(s,t) = \|\mathbf{W}_1\|^p_{p-var;[s,t]} + \|\mathbf{W}_2\|^p_{p-var;[s,t]} + \|\gamma^\mu\|_{1-var;[s,t]} + \|\gamma^\nu\|_{1-var;[s,t]}.$$

We have

$$\begin{aligned}&\rho_{\alpha-\text{Höl};[0,T]}\Big(\Theta_{b,\sigma}(\mu,\xi,\mathbf{W}_1), \Theta_{b,\sigma}(\nu,\chi,\mathbf{W}_2)\Big)\\ &\qquad\le C\Big(|\xi-\chi| + \rho_{\alpha-\text{Höl};[0,T]}(\mathbf{W}_1,\mathbf{W}_2) + \rho_{1-\text{Höl};[0,T]}(\gamma^\mu,\gamma^\nu)\Big)\exp\Big(\mathbf{M}_{\beta,[0,T]}(\omega)\Big).\end{aligned}$$

Proposition 13.4.1 shows continuity in measure pointwise for each geometric rough path $\mathbf{W}$. Therefore, to prove joint continuity via Moore-Osgood we verify the uniform continuity condition.

Let $\mu_k, \mu \in \mathcal{P}_2\big(G\Omega_\alpha(\mathbb{R}^d)\big)$ and $\mathbb{W}^{(2)}_{\rho_{\alpha-\text{Höl};[0,T]}}(\mu_k,\mu) \to 0$. Then we also have

$$\lim_{k\to\infty}\|\gamma^{\mu_k} - \gamma^\mu\|_{1-var;[0,T]} = 0.$$

Hence there must exist an $C' \in \mathbb{N}$ such that

$$\sup_{k>C'}\|\gamma^{\mu_k}\|_{1-var;[0,T]} \le \|\gamma^\mu\|_{1-var;[0,T]} + 1.$$

Similarly, by choosing C' large enough

$$\sup_{k>C'} \|\mathbf{W}_k\|_{p-var;[0,T]}^p \leq \Big(\|\mathbf{W}\|_{p-var;[0,T]} + 1\Big)^p.$$

Thus

$$\begin{aligned}
\sup_{k>C'} \rho_{\alpha-\text{Höl};[0,T]}&\Big(\Theta_{b,\sigma}(\mu_k, \xi_k, \mathbf{W}_1), \Theta_{b,\sigma}(\mu_k, \xi_k, \mathbf{W}_2)\Big) \\
&\leq C\rho_{\alpha-\text{Höl};[0,T]}(\mathbf{W}_1, \mathbf{W}_2) \\
&\quad \cdot \exp\Big(\mathbf{M}_{\beta,p}(\mathbf{W}_1) + \mathbf{M}_{\beta,p}(\mathbf{W}_2)\Big) \exp\Big(\|\gamma^\mu\|_{1-var;[0,T]} + 1\Big), \\
\sup_{k>C'} \rho_{\alpha-\text{Höl};[0,T]}&\Big(\Theta_{b,\sigma}(\mu_k, \xi, \mathbf{W}_k), \Theta_{b,\sigma}(\mu_k, \chi, \mathbf{W}_k)\Big) \\
&\leq C|\xi - \chi| \cdot \exp\Big(\big(\|\mathbf{W}\|_{p-\text{var};[0,T]} + 1\big)^p\Big) \exp\Big(\|\gamma^\mu\|_{1-var;[0,T]} + 1\Big), \\
\sup_{k>C'} \rho_{\alpha-\text{Höl};[0,T]}&\Big(\Theta_{b,\sigma}(\mu, \xi_k, \mathbf{W}_k), \Theta_{b,\sigma}(\nu, \xi_k, \mathbf{W}_k)\Big) \\
&\leq C\rho_{1-\text{Höl};[0,T]}(\gamma^\mu, \gamma^\nu) \\
&\quad \cdot \exp\Big(\big(\|\mathbf{W}\|_{p-\text{var};[0,T]} + 1\big)^p\Big) \exp\Big(\|\gamma^\mu\|_{1-var;[0,T]} + \|\gamma^\nu\|_{1-var;[0,T]}\Big)
\end{aligned}$$

which implies uniform continuity. $\square$

Chapter 14

Support Theorem

We state and prove representations of the support of McKean-Vlasov Equations in terms of the particle systems associated to the quantizations that we constructed in Chapter 12. We introduce a collection of sets of paths that to the best of our knowledge have not previously been described in another work. These sets are all subsets of $C^{\alpha,0}([0,T];\mathbb{R}^d)$ and are defined solely with respect to the RKHS $\mathcal{H}$, the Hölder norm $\|\cdot\|_\alpha$ and the coefficients of the Rough Differential Equation (13.0.1).

In order to provide a clear exposition of the construction of the support, we briefly summarise the upcoming subsections: from the previous chapter we have obtained a sequence of quantizations $\mathbf{q}_n$ for the the law of the enhanced Brownian motion with codebooks $\mathbf{C}_n$.

- For each quantization, we solve the system of interacting ODEs in Section 14.1.1 (see Equation (14.1.2)) by replacing the path of Brownian motion by the associated codebook path and replacing the law of the Brownian motion by the quantization
- By associating to each of these ODEs the probability weight associated to the codebook element driving the equation, we obtain a finite support measure in Section 14.1.2 (see Equation (14.1.3)). We call this the quantization of the McKean-Vlasov Equation. This sequence of finite support measures converges to the law of the McKean-Vlasov Equation.
- In Section 14.2.1, for fixed n, we replace the law of the McKean-Vlasov Equation inside the canonical skeleton process by the quantization of the McKean-Vlasov Equation (see Definition 14.2.2). These paths will not generally be contained in the support of the McKean-Vlasov Equation. However, a ball of large enough radius will have positive measure (see Lemma 14.2.3).
- In Section 14.2.2, we show that for n chosen large enough, an ε ball around this collection of paths will be a closed set of measure 1. By taking an intersection of these sets, we show the set of limit points has measure 1 (see Theorem 14.2.4).
- Finally in Section 14.3, we extend our work to the case where the McKean-Vlasov Equation has a random initial condition (see Theorem 14.3.6).
- An example is presented in Section 14.4.

The results of this chapter can be found in the preprint [CRS19, Section 5].

14.1 The Skeleton Process for a McKean-Vlasov Equations

The law of a McKean-Vlasov equation is deterministic; it is not dependent on the choice of driving noise. The Occupation Measure path is of bounded variation and does not interact with the noise. Thus when the Occupation Measure path is known, McKean-Vlasov Equations can be thought of as classical Rough Differential Measure Equations with a drift term. Thus, we can define a skeleton process in the following classical sense:

Definition 14.1.1. *Let $\mathcal{L}^{\mathbf{W}}$ be the law of an enhanced Brownian motion. Let b and σ satisfy Assumption 13.2.2. Let $\xi \in \mathbb{R}^d$. Let $\mathcal{L}^{\mathbf{X}}$ be the unique fixed point of the operator $\Psi_{\mathcal{L}^{\mathbf{W}}}$. Then we define the* True Skeleton Operator $\mathbf{\Phi}' : \mathcal{H} \times \mathbb{R}^d \to G\Omega_\alpha(\mathbb{R}^d)$ *to be the operator that maps the element of the RKHS to the solution of the ODE*

$$d\mathbf{\Phi}'(h,\xi)_t = b(\mathbf{\Phi}'(h,\xi)_t)d\gamma_t^{\mathcal{L}^{\mathbf{X}}} + \sigma(\mathbf{\Phi}'(h,\xi)_t)dh_t, \quad \mathbf{\Phi}(h,\xi)_0 = \xi. \tag{14.1.1}$$

It is important to emphasise that the *true skeleton operator* (14.1.1) is dependent on the measure $\mathcal{L}^{\mathbf{X}}$ and as such it cannot be solved without knowing the law exogenously. The main contribution of this Section is how one navigates around this issue.

14.1.1 Interacting Particle system derived from Quantization

We introduce a system of interacting Ordinary Differential Equations that model the dynamics of the McKean-Vlasov Equation.

Definition 14.1.2. *Let $\xi \in \mathbb{R}^d$. Let $\mathcal{L} \in \mathcal{P}_c(G\Omega_\alpha(\mathbb{R}^{d'}))$ be a finitely supported measure over the space of geometric rough paths with the form $\mathcal{L} = \sum_{j=1}^n \mathfrak{p}_j \delta_{\mathbf{W}_j}$ where $(\mathfrak{p}_j)_{j=1,\dots,n}$ is a probability vector. For a codebook $\mathbf{C} := \{\mathbf{W}_j : j=1,...,n\}$, let $\mathbf{W} := \oplus_{j=1}^n \mathbf{W}_j$ and let $\tilde{\mathbf{W}}$ be the extension of $\mathbf{W}$ to $G^M\big(\mathbb{R}^{d'\times n}\big)$ where M is the largest integer such that $M\alpha < 1$. Let b and σ satisfy Assumption 13.2.2. Let B and Σ be as in Definition 13.2.5.*

*Then we define the $\mathcal{L}$-*Interacting Particle System *to be the solution to the Rough Differential Equation*

$$d\mathbf{\Phi}(\mathcal{L})_t = B\big(\mathbf{\Phi}(\mathcal{L})_t\big)dt + \Sigma\big(\mathbf{\Phi}(\mathcal{L})_t\big)d\tilde{\mathbf{W}}_t, \quad \mathbf{\Phi}(\mathcal{L})_0 = \oplus_{j=1}^n \xi \in \mathbb{R}^{d\times n} \tag{14.1.2}$$

taking values in $G\Omega_\alpha\Big(\mathbb{R}^{d\times n}\Big)$.

An important detail about this object is that this is a *finite dimensional* system of Rough Differential Equations. This system of interacting equations can be solved without having to consider any measures.

The existence and uniqueness of the ODE (14.1.2) is standard. In particular, by Theorem 13.2.8 the solution to Equation (14.1.2) is independent of the choice of $\tilde{\mathbf{W}}$ and only on $\mathbf{W}$.

14.1.2 Quantization of the McKean-Vlasov Equation

We use the interacting particle system (14.1.2) to obtain a law that approximates the law of the McKean-Vlasov Equation (13.0.1).

Definition 14.1.3. *Let $\mathcal{L} \in \mathcal{P}_c(G\Omega_\alpha(\mathbb{R}^{d'}))$ be a finite support measure over the space of geometric rough paths with the form $\mathcal{L} = \sum_{m=1}^n \mathfrak{p}_m \delta_{\mathbf{W}^m}$ where $(\mathfrak{p}_m)_{m=1,\dots,n}$ is a probability vector. Let b and σ satisfy Assumption 13.2.2.*

Let $\mathbf{\Phi}(\mathcal{L})$ be the solution to Equation (14.1.2). *Let $\pi^{(m)} : G\Omega_\alpha(\mathbb{R}^{d\times n}) \to G\Omega_\alpha(\mathbb{R}^d)$ be the quotient operator obtained by extending the projection $\langle \cdot, e_{(\cdot,m)}\rangle$. Then we define the* Law of the $\mathcal{L}$-Interacting Particle System *to be the finite measure over $G\Omega_\alpha(\mathbb{R}^d)$*

$$\mathcal{L}^{\mathbf{\Phi}(\mathcal{L})} := \sum_{m=1}^n \mathfrak{p}_m \delta_{\pi^{(m)}[\mathbf{\Phi}(\mathcal{L})]}. \tag{14.1.3}$$

Substituting a quantization of the Brownian motion into an Interacting Particle System and taking its law, we obtain a quantization for the McKean-Vlasov Equation.

Proposition 14.1.4. *Let $\mathcal{L}^{\mathbf{W}}$ be the law of enhanced Brownian motion. Let $\mathcal{L}^{\mathbf{W}} \circ \mathbf{q}_n^{-1}$ be the sequence of quantizations of the enhanced Brownian motion from Definition 12.2.6.*

Let $\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})}$ be the sequence of quantizations for the McKean-Vlasov obtained from the sequence of finite support measures $\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1}$. Then $\Xi\Big[\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1}\Big]=\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})}$ so that

$$\mathbb{W}^{(2)}_{\rho_{\alpha-H\ddot{o}l;[0,T]}}\Big(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\mathcal{L}^{\mathbf{X}}\Big)\lesssim\Big(\log(n)\Big)^{\alpha-1/2}.$$

Proof. We have $\Xi\Big[\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1}\Big]=\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})}$ and $\Xi\Big[\mathcal{L}^{\mathbf{W}}\Big]=\mathcal{L}^{\mathbf{X}}$. By Proposition 13.3.2, we have

$$\mathbb{W}^{(2)}_{\rho_{\alpha-\text{Höl};[0,T]}}\Big(\Xi\big[\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1}\big],\Xi\big[\mathcal{L}^{\mathbf{W}}\big]\Big)\lesssim\mathbb{W}^{(2)}_{\rho_{\alpha-\text{Höl};[0,T]}}\Big(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1},\mathcal{L}^{\mathbf{W}}\Big).$$

Apply Theorem 12.2.8 for the rate of convergence. □

14.2 The Support of the McKean-Vlasov Equation

The following result immediately holds from the methods laid out in [FV10b, Chapter 19].

Theorem 14.2.1. *Let $\mathcal{L}^{\mathbf{W}}$ be the law of an enhanced Brownian motion. Let $\xi\in\mathbb{R}^d$. Let b and σ satisfy Assumption 13.2.2. Let $\mathcal{L}^{\mathbf{X}}$ be the law of the McKean-Vlasov Equation* (13.0.1). *Then the support of $\mathcal{L}^{\mathbf{X}}$ can be characterised with respect to the rough path Hölder metric by*

$$\operatorname{supp}(\mathcal{L}^{\mathbf{X}})=\overline{\Big\{\mathbf{\Phi}'(h,\xi):h\in\mathcal{H}\Big\}}^{\rho_{\alpha-H\ddot{o}l;[0,T]}}\tag{14.2.1}$$

where $\mathbf{\Phi}'$ is the true skeleton operator from Definition 14.1.1.

This is not a meaningful result as the true skeleton operator includes a priori knowledge of the law of the McKean-Vlasov Equation. This measure can be proved to exist, but constructing it is another matter. We overcome this issue via functional quantization.

14.2.1 Quantized Skeleton of McKean-Vlasov Equation

We use the quantized McKean-Vlasov Equation to construct a skeleton process that approximates the true skeleton process.

Definition 14.2.2. *Let $\mathcal{L}^{\mathbf{W}}$ be the law of an enhanced Brownian Motion. Let $\mathbf{q}_n$ be the sequence of quantizations of $\mathcal{L}^{\mathbf{W}}$ constructed in Definition 12.2.6. Let $\xi\in\mathbb{R}^d$ and let $h\in\mathcal{H}$ and denote $\mathbf{h}=S_2[h]$. Let b and σ satisfy Assumption 13.2.2.*

Fix $\varepsilon>0$ and choose $n\in\mathbb{N}$ such that

$$\rho_{\alpha-H\ddot{o}l;[0,T]}\Big(\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\xi,\mathbf{h}),\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{X}},\xi,\mathbf{h})\Big)\le\varepsilon,\tag{14.2.2}$$

and we define the sets $A_\varepsilon(\mathbf{h})$ as

$$A_\varepsilon(\mathbf{h}):=\Big\{\mathbf{Y}\in G\Omega_\alpha(\mathbb{R}^d):\rho_{\alpha-H\ddot{o}l;[0,T]}\Big(\mathbf{Y},\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\xi,\mathbf{h})\Big)<\varepsilon\Big\}.\tag{14.2.3}$$

We emphasise that the choice of n will not be uniform over all choices of $h\in\mathcal{H}$. Also note that $\mathbf{\Phi}'(h,\xi)=\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{X}},\xi,\mathbf{h})$. The first goal is to show that each of these sets contains an element of the $\operatorname{supp}(\mathcal{L}^{\mathbf{W}})$, regardless of ε.

Lemma 14.2.3. *Let $h\in\mathcal{H}$ and $\mathbf{h}=S_2[h]$. Then $\forall\varepsilon>0$, the open sets $A_\varepsilon(\mathbf{h})$ of Definition 14.2.2 have positive measure with respect to $\mathcal{L}^{\mathbf{X}}$,*

$$\mathcal{L}^{\mathbf{W}}\Big[A_\varepsilon(\mathbf{h})\Big]>0.$$

Proof. The condition for $A_\varepsilon(\mathbf{h})$ in Equation (14.2.2) is the key. It ensures that for any choice of $\varepsilon>0$, we have $\mathbf{\Phi}'(h,\xi)\in A_\varepsilon(\mathbf{h})$. By Theorem 14.2.1, we have that any open set $B\subseteq G\Omega_\alpha(\mathbb{R}^d)$

containing a path $\mathbf{\Phi}'(h,\xi)$ and for any choice of $h\in\mathcal{H}$, we have

$$\mathcal{L}^{\mathbf{X}}[B]>0.$$

$\square$

14.2.2 The Support of McKean-Vlasov Equations

We now formulate our statement of the support theorem of McKean-Vlasov Equations:

Theorem 14.2.4. *Let $\mathcal{L}^{\mathbf{W}}$ be the law of an enhanced Brownian motion. Let $\mathbf{q}_n$ be the sequence of quantizations obtained in Definition 12.2.6. Let $\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})}$ be the law of the Interacting Particle System driven by the quantization constructed in Definition 14.1.3. Let $\xi\in\mathbb{R}^d$. Suppose that b and σ satisfy Assumption 13.2.2. Then the law of the solution to the McKean-Vlasov Equation* (13.0.1) *satisfies*

$$\operatorname{supp}(\mathcal{L}^{\mathbf{X}})=\bigcap_{m=1}^{\infty}\overline{\bigcup_{n\geq m}\Big\{\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\xi,\mathbf{h}):h\in\mathcal{H},\mathbf{h}=S_2(h)\Big\}}^{\rho_{\alpha-\text{Höl};[0,T]}}. \tag{14.2.4}$$

We emphasise that this expression of the support is only dependent on:

- The RKHS of Brownian motion $\mathcal{H}$ and the initial condition $\xi\in\mathbb{R}^d$
- The coefficients b and σ
- The sequence of Systems of Interacting Particles $\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})$ which is in turn dependent on
  - The coefficients b and σ
  - The sequence of quantizations $\mathbf{q}_n$ which are only dependent on $\mathcal{H}$ and $\|\cdot\|_\alpha$.

We have not solved the law of the McKean-Vlasov Equation or the Occupation measure path at any point of this approach.

Proof. For the simplicity of the proof, we rely on Theorem 14.2.1 for an expression of $\operatorname{supp}(\mathcal{L}^{\mathbf{X}})$. By Proposition 14.1.4, we have that the law of the Interacting Particle System converges to the law of the McKean-Vlasov Equation as $n\to\infty$. Fix $h\in\mathcal{H}$ and $m\in\mathbb{N}$. Then $\forall l\geq m$

$$\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_l^{-1})},\xi,\mathbf{h})\in\overline{\bigcup_{n\geq m}\Big\{\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\xi,\mathbf{h}):h\in\mathcal{H},\mathbf{h}=S_2(h)\Big\}}^{\rho_{\alpha-\text{Höl};[0,T]}}.$$

Since this is closed, we have that the limit of these paths is also contained so

$$\mathbf{\Phi}'(h,\xi)=\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{X}},\xi,\mathbf{h})\in\overline{\bigcup_{n\geq m}\Big\{\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\xi,\mathbf{h}):h\in\mathcal{H},\mathbf{h}=S_2(h)\Big\}}^{\rho_{\alpha-\text{Höl};[0,T]}}. \tag{14.2.5}$$

Finally, Equation (14.2.5) holds for any choice of $m\in\mathbb{N}$, so it must be contained in the intersection over all m. This was true for any choice of $h\in\mathcal{H}$, so it is also true for all $h\in\mathcal{H}$. Thus

$$\Big\{\mathbf{\Phi}'(h,\xi):h\in\mathcal{H}\Big\}\subset\bigcap_{m=1}^{\infty}\overline{\bigcup_{n\geq m}\Big\{\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\xi,\mathbf{h}):h\in\mathcal{H}\Big\}}^{\rho_{\alpha-\text{Höl};[0,T]}}.$$

Finally, as the right hand side is closed, we can take a closure on the left hand side to achieve the first implication.

Now we show the reverse implication. Suppose $\mathbf{Y}\in G\Omega_\alpha(\mathbb{R}^d)$ such that

$$\mathbf{Y}\in\bigcap_{m=1}^{\infty}\overline{\bigcup_{n\geq m}\Big\{\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_n^{-1})},\xi,\mathbf{h}):h\in\mathcal{H}\Big\}}^{\rho_{\alpha-\text{Höl};[0,T]}}.$$

Then there must exist a subsequence n_k and a sequence of $h_k \in \mathcal{H}$ such that

$$\lim_{k\to\infty} \rho_{\alpha-\text{Höl};[0,T]}\Big(\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_{n_k}^{-1})}, \xi, \mathbf{h}_k), \mathbf{Y}\Big) = 0.$$

Further, we know the sequence satisfies $\lim_{k\to\infty} n_k = \infty$, since $\mathbf{Y}$ is in the intersect over all $m \in \mathbb{N}$. Thus the weak limit of $\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}}\circ\mathbf{q}_{n_k}^{-1})}$ must just be $\mathcal{L}^{\mathbf{X}}$ as $k \to \infty$.

By Theorem 13.4.2, we have joint continuity of $\Theta_{b,\sigma}$. Therefore, taking the limit in the measure variable first, we get

$$\lim_{k\to\infty} \rho_{\alpha-\text{Höl};[0,T]}\Big(\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{X}}, \xi, \mathbf{h}_k), \mathbf{Y}\Big) = \lim_{k\to\infty} \rho_{\alpha-\text{Höl};[0,T]}\Big(\mathbf{\Phi}'(h_k, \xi), \mathbf{Y}\Big) = 0,$$

which just means that $\mathbf{Y} \in \overline{\{\mathbf{\Phi}'(\mathbf{h}, \xi) : h \in \mathcal{H}\}}^{\rho_{\alpha-\text{Höl};[0,T]}}$. $\square$

14.3 Random Initial Conditions

An apparent limitation of the previous Section is that we restrict ourselves to McKean-Vlasov Equations with constant initial conditions. However, there is an easy extension to the case where the initial condition is random.

We introduce a Theorem first proved in [CFN97] that allows for the consideration of random initial conditions.

Theorem 14.3.1 ([CFN97])**.** *Let $F : \Omega \times \mathbb{R}^d \to E$ be a random variable taking values in a Banach space E such that $x \mapsto F(\omega, x)$ is continuous for each ω. Suppose that $G : \mathcal{H} \times \mathbb{R}^d \to E$ is a uniform skeleton of F. Suppose that ζ is an d-dimensional random variable with skeleton ϕ. Then $\tilde{G}(h) := G(h, \phi(h))$ is a skeleton of $\tilde{F}(\omega) := F(\omega, \zeta(\omega))$.*

We now turn to the McKean-Vlasov Equation

$$dX_t = \sigma(X_t)d\mathbf{W}_t + b(X_t)d\gamma^{\mathcal{L}_t^{\mathbf{X}}}, \quad X_0 \sim \xi \in \mathcal{P}_r(\mathbb{R}^d) \tag{14.3.1}$$

where $r > 1$.

Following in the footsteps of Definition 12.2.6, we construct a quantization for the law $\xi \times \mathcal{L}^{\mathbf{W}}$ over $\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^{d'})$.

Definition 14.3.2. *Let $r > 1$. Let $\mathcal{L}^W$ be the law of a Brownian motion over $C^{\alpha,0}([0,T]; \mathbb{R}^{d'})$. Let $\xi \in \mathcal{P}_r(\mathbb{R}^d)$. Let $m, n \in \mathbb{N}$.*

1. *By Theorem B.1.2, there exists a codebook $\mathfrak{C}_m^{(1)} \subset \mathbb{R}^d$ that is an m-stationary set with Voronoi partition $\mathfrak{S}_m^{(1)}$. Let $\mathbf{C}_n^{(2)}$ be the n element codebook constructed in Definition 12.2.6 with partition $\mathbf{S}_n^{(2)}$.*

2. *Let $\mathbf{C}_{m,n} := \mathfrak{C}_m^{(1)} \times \mathbf{C}_n^{(2)}$ be a sequence of codebooks over $\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^{d'})$ and let $\mathbf{S}_{m,n} := \mathfrak{S}_m^{(1)} \times \mathbf{S}_n^{(2)}$ be a partition of $\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^{d'})$. Let $\mathbf{q}_{m,n}$ be the Quantization with codebook $\mathbf{C}_{m,n}$ and partition $\mathbf{S}_{m,n}$. Then $|\mathbf{C}_{m,n}| = m \cdot n$.*

3. *By combining Equation* (B.2.1) *and Theorem 12.2.8, the rate of convergence is*

$$\tfrac{1}{m^{1/d}} + \Big(\log(n)\Big)^{\alpha-1/2} \approx \left(\int_{\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^{d'})} d_{|\cdot|\times\rho_{\alpha\text{-Höl}}}\Big((x,\mathbf{Y}), \mathbf{q}_{m,n}(x,\mathbf{Y})\Big)^r d[\xi \times \mathcal{L}^{\mathbf{W}}](x,\mathbf{Y}) \right)^{1/r}. \tag{14.3.2}$$

4. *By choosing $m \approx [\log(n)]^{(1/2-\alpha)d}$ and rescaling, we obtain the sequence of quantizations*

$$\left(\int_{\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^{d'})} d_{|\cdot|\times\rho_{\alpha\text{-Höl}}}\Big((x,\mathbf{Y}), \mathbf{q}_n(x,\mathbf{Y})\Big)^r d[\xi\times\mathcal{L}^{\mathbf{W}}](x,\mathbf{Y}) \right)^{1/r}$$
$$\approx \left[\log\left(\tfrac{n}{[(1/2-\alpha)d]^{(1/2-\alpha)d}}\right) - \log\left(\mathcal{W}\left(\frac{n^{\frac{1}{(1/2-\alpha)d}}}{(1/2-\alpha)d}\right)^{(1/2-\alpha)d}\right)\right]^{(\alpha-1/2)d} \tag{14.3.3}$$

where, as in Proposition 12.2.4, $\mathcal{W}$ is the Lambert W function.

Next, following Definition 14.1.3, we define a new interacting particle system.

Definition 14.3.3. *Let $\mathcal{L} \in \mathcal{P}_c(\mathbb{R}^d \times G\Omega_\alpha(\mathbb{R}^{d'}))$ be a finite support measure of the form $\mathcal{L} = \sum_{j=1}^n \mathfrak{p}_j \delta_{(x_j,\mathbf{W}_j)}$ where $(\mathfrak{p}_j)_{j=1,\dots,n}$ is a probability vector. For codebook $\mathbf{C} := \{(x_j,\mathbf{W}_j) : j = 1,\dots,n\}$, let $\mathbf{W} := \oplus_{j=1}^n \mathbf{W}_j$ and $X = \oplus_{j=1}^n x_j \in \mathbb{R}^{d\times n}$. Let $\tilde{\mathbf{W}}$ be the lift of the path $\mathbf{W}$ to a rough path. Let b and σ satisfy Assumption 13.2.2. Let B and Σ be as in Definition 13.2.5.*

*Then we define the $\mathcal{L}$-*Interacting Particle System *with random initial condition to be the solution to the Rough Differential Equation*

$$d\mathbf{\Phi}(\mathcal{L})_t = B\big(\mathbf{\Phi}(\mathcal{L})_t\big)dt + \Sigma\big(\mathbf{\Phi}(\mathcal{L})_t\big)d\tilde{\mathbf{W}}_t, \quad \mathbf{\Phi}(\mathcal{L})_0 = X. \tag{14.3.4}$$

We also define the law of the $\mathcal{L}$ Interacting Particle system in $\mathcal{P}_c\Big(G\Omega_\alpha(\mathbb{R}^d)\Big)$ to be

$$\mathcal{L}^{\mathbf{\Phi}(\mathcal{L})} := \sum_{m=1}^n \mathfrak{p}_m \delta_{\pi^{(m)}[\mathbf{\Phi}(\mathcal{L})]}.$$

As with Theorem 13.2.8, the paths of this law are dependent only on $\mathbf{W}$ and not of the lift of $\tilde{\mathbf{W}}$. In this definition we do not limit ourselves to the case where many of the x_j values are repeated. We use the quantization of the measure $\xi\times\mathcal{L}^{\mathbf{W}}$ constructed in Definition 14.3.2 to solve the law of an Interacting Particle system that approximates the true law of the McKean-Vlasov Equation

Proposition 14.3.4. *Let $\mathcal{L}^{\mathbf{W}}$ be the law of the enhanced Brownian motion. Let $[\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1}$ be the sequence of quantizations of the enhanced Brownian motion from Definition 14.3.2.*

Let $\mathcal{L}^{\mathbf{\Phi}([\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1})}$ be the sequence of quantizations for the McKean-Vlasov obtained from the sequence of finite support measures $[\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1}$.

Then $\Xi\Big[[\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1}\Big] = \mathcal{L}^{\mathbf{\Phi}([\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1})}$ so that

$$\mathbb{W}^{(1)}_{\rho_{\alpha-Höl}}\Big(\mathcal{L}^{\mathbf{\Phi}([\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1})}, [\xi\times\mathcal{L}^{\mathbf{X}}]\Big)$$
$$\lesssim \left[\log\left(\tfrac{n}{[(1/2-\alpha)d]^{(1/2-\alpha)d}}\right) - \log\left(\mathcal{W}\left(\frac{n^{\frac{1}{(1/2-\alpha)d}}}{(1/2-\alpha)d}\right)^{(1/2-\alpha)d}\right)\right]^{(\alpha-1/2)d}.$$

Proof. Same method as Proposition 14.1.4 with Equation (14.3.3). □

14.3.1 Statement for the Support

Using classical tools, we combine the results of Theorem 14.2.1 with [CFN97] for this next Theorem:

Theorem 14.3.5. *Let $r > 1$. Let $\xi \in P_r(\mathbb{R}^e)$. Let $\mathcal{L}^{\mathbf{W}}$ be the law on an enhanced Brownian motion. Let b and σ be defined in Definition 13.2.9. Suppose that Assumption 13.2.10 is satisfied. Let $\mathcal{L}^{\mathbf{X}}$ be the law of the McKean-Vlasov Equation* (14.3.1)*. Then the support of $\mathcal{L}^{\mathbf{X}}$ can be characterised with respect to the Rough Path Hölder metric by*

$$\mathrm{supp}(\mathcal{L}^{\mathbf{X}}) = \overline{\Big\{ \mathbf{\Phi}'(h,x) : h\in\mathcal{H}, x\in\mathrm{supp}(\xi)\Big\}}^{\rho_{\alpha-Höl;[0,T]}}$$

where $\mathbf{\Phi}'$ is the true skeleton operator from Definition 14.1.1.

This Theorem is not meaningful and our final Theorem is the culmination of this work:

Theorem 14.3.6. *Let $r > 1$. Let $\xi \in \mathcal{P}_r(\mathbb{R}^d)$. Let $\mathcal{L}^{\mathbf{W}}$ be the law of an enhanced Brownian motion. Let $\mathbf{q}_n$ be the sequence of quantizations obtained in Definition 14.3.2. Let $\mathcal{L}^{\mathbf{\Phi}([\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1})}$ be the law of the Interacting Particle System driven by the quantization constructed in Definition 14.3.3. Suppose that b and σ satisfy Assumption 13.2.2. Then the law of the solution to the McKean-Vlasov Equation* (13.0.1) *satisfies*

$$\operatorname{supp}(\mathcal{L}^{\mathbf{X}}) = \bigcap_{m=1}^{\infty} \overline{\bigcup_{n\geq m} \Big\{\Theta_{b,\sigma}(\mathcal{L}^{\mathbf{\Phi}([\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1})}, x, \mathbf{h}) : h\in\mathcal{H}, x\in\operatorname{supp}(\xi)\Big\}}^{\rho_{\alpha-H\ddot{o}l;[0,T]}}. \quad (14.3.5)$$

Proof. See proof of Theorem 14.2.4 with Proposition 14.3.4 and Theorem 13.4.2. □

14.4 Example

Firstly, to illustrate a relevant example in the simplest framework possible, we consider an SDE where the diffusion is constant. Thus there is no need to consider rough paths and the enhancement of the rough quantization and we rely only on pathspace quantization.

Example 14.4.1. *Let $x_0 \in \mathbb{R}^d$. Let W be a d-dimensional Brownian motion. Consider the diffusion*

$$X_t = x_0 + \int_0^t b(X_s, \mathcal{L}_s^X) ds + W_t \quad (14.4.1)$$

where b is adequately regular in spacial and measure derivatives.

For a quantization q_n of $\mathcal{L}^W$ with codebook $\mathfrak{C}_n$ and partition $\mathfrak{S}_n$, we obtain the probability vector $(\mathfrak{p}^i)$ such that $\mathfrak{p}^i = \mathcal{L}[\mathfrak{s}^i]$. For each $\mathfrak{h}^i \in \mathfrak{C}_n$, we solve the deterministic system of equations

$$\Phi^n[\mathfrak{h}^i]_t = x_0 + \int_0^t b\Big(\Phi^n[\mathfrak{h}^i]_s, \sum_{\mathfrak{h}^i\in\mathfrak{C}_n} \mathfrak{p}^i \cdot \delta_{\Phi^n[\mathfrak{h}^i]_s}\Big) ds + \mathfrak{h}^i_t$$

Then

$$\mathcal{L}_n^X := \sum_{\mathfrak{h}^i\in\mathfrak{C}_n} \mathfrak{p}^i \cdot \delta_{\Phi^n[\mathfrak{h}^i]} \to \mathcal{L}^X$$

strongly over pathspace.

Thus for any $h \in \mathcal{H}$, the dynamics of the ordinary differential equation

$$\Phi[h, \mathcal{L}_n^X]_t = x_0 + \int_0^t b\Big(\Phi[h,\mathcal{L}_n^X]_s, (\mathcal{L}_n^X)_s\Big) ds + h_t$$

will converge as $n \to \infty$ to the dynamics of the true skeleton process

$$\Phi[h]_t = x_0 + \int_0^t b\Big(\Phi[h]_s, \mathcal{L}_s^X\Big) ds + h_t.$$

Thus we are able to write the support of the McKean-Vlasov equation (14.4.1) *as*

$$\operatorname{supp}(\mathcal{L}^X) = \bigcap_{m=1}^{\infty} \overline{\bigcup_{n\geq m} \Big\{\Phi[h,\mathcal{L}_n^X], h\in\mathcal{H}\Big\}}^{\alpha}.$$

Next, we wish to demonstrate the support of a diffusion where the measure dependency determines the nature of the support.

Example 14.4.2 (Delayed Non-degenerate Affine linear SDE)**.** *Consider the 3-dimensional*

$$\begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \Big(\mathbb{E}\Big[|X(t)|\Big] - 1\Big) \wedge 0 \end{pmatrix} + \begin{pmatrix} X_1(t), & 0, & 0 \\ 0, & X_2(t), & 0 \\ 0, & 0, & X_3(t) \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \\ dW_3(t) \end{pmatrix}, \quad \begin{pmatrix} X_1(0) \\ X_2(0) \\ X_3(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

By following [CF10] *with addition of a drift term which is accounted for using the methods of* [Nua06, Theorem 2.3.2]*, we can see that on a small time interval Hörmander's condition is only satisfied when we restrict to $(X_1(t), X_2(t)) \in \mathbb{R}^2$. It is not satisfied over $\mathbb{R}^3$ locally around $t = 0$. However, once $\mathbb{E}[|X(t)| > 1$ Hörmander's condition is satisfied and a density exists. Therefore, the value $t' > 0$ such that $\mathbb{E}[|X(t')|] = 1$ is key to the support of this McKean-Vlasov Equation.*

Given our sequence of quantizations, we obtain a sequence of laws for the associated particle systems. For these, we compute the first t'_n that satisfies

$$\mathbb{E}\Big[\Big|\Theta(\mathcal{L}^{\mathbf{W}} \circ \mathbf{q}_n^{-1}, \mathbf{W})_{t'_n}\Big|\Big] = 1$$

On the interval $[0, t' \wedge t'_n]$ we have that

$$\overline{\Big\{\Theta(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}} \circ \mathbf{q}_n^{-1})}, \xi, \mathbf{h}) : h \in \mathcal{H}\Big\}}^{\rho_{\alpha-Höl;[0,t'\wedge t'_n]}} = \overline{\Big\{\Theta(\mathcal{L}^{\mathbf{X}}, \xi, \mathbf{h}) : h \in \mathcal{H}\Big\}}^{\rho_{\alpha-Höl;[0,t'\wedge t'_n]}}.$$

Similarly, we also have that on the interval $[t' \vee t'_n, T \vee t' \vee t'_n]$

$$\overline{\Big\{\Theta(\mathcal{L}^{\mathbf{\Phi}(\mathcal{L}^{\mathbf{W}} \circ \mathbf{q}_n^{-1})}, \xi, \mathbf{h}) : h \in \mathcal{H}\Big\}}^{\rho_{\alpha-Höl;[t'\vee t'_n, T\vee t'\vee t'_n]}} = \overline{\Big\{\Theta(\mathcal{L}^{\mathbf{X}}, \xi, \mathbf{h}) : h \in \mathcal{H}\Big\}}^{\rho_{\alpha-Höl;[t'\vee t'_n, T\vee t'\vee t'_n]}}.$$

The sets are not equal outside these intervals. However, as $n \to \infty$ we have that the deterministic $t'_n \to t'$. Thus taking the limit points of these sets as in (14.2.4) *provides the support for the entire interval.*

Part V

Small Ball Probabilities for Gaussian rough paths

Chapter 15

Enhanced Gaussian Inequalities

In this Chapter, we prove a series of inequalities for Gaussian measures that we will use when proving the small ball probability results of Chapter 16. These results can be found in the preprint [Sal20, Section 3].

We denote by $\mathbb{B}_\alpha(\mathbf{h},\varepsilon) := \{\mathbf{X} \in G\Omega_\alpha(\mathbb{R}^d) : d_\alpha(\mathbf{h},\mathbf{X}) < \varepsilon\}$ the unit ball in the metric space of rough paths with respect to the homogeneous α-Hölder norm.

15.1 Translation Inequalities

When working with measures over Euclidean space, one is able to translate results relating to small ball probabilities using the density of the measure. For Gaussian measures, the density takes a particular form and so the probability of a small ball centred at a point can be calculated using only the small ball probabilities centred around 0.

When working on Banach spaces, this property holds with respect to translations of Cameron Martin space paths as any Cameron Martin translation maps the Gaussian measure to another Gaussian measure that is absolutely continuous with respect to the original. In this first section, we extend these results to enhanced Gaussian measures using the rough path translation operator (see Definition C.2.10).

Lemma 15.1.1 (Anderson's Inequality for Gaussian rough paths)**.** *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 and let $\mathcal{L}^{\mathbf{W}}$ be the law of the enhanced Gaussian. Then $\forall \mathbf{X} \in G\Omega_\alpha(\mathbb{R}^d)$*

$$\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{X},\varepsilon)\Big] \leq \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1},\varepsilon)\Big]. \tag{15.1.1}$$

Proof. See for instance [Lif13]. □

Lemma 15.1.2 (Cameron Martin rough path formula)**.** *Let $\mathcal{L}^W$ be a Gaussian measure and let $\mathcal{L}^{\mathbf{W}}$ be the law of the lift to the Gaussian rough path. Let $h \in \mathcal{H}$ and denote $\mathbf{h} = S_2[h]$. Then*

$$\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{h},\varepsilon)\Big] \geq \exp\left(\frac{-\|h\|_{\mathcal{H}}^2}{2}\right)\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1},\varepsilon)\Big]$$

Proof. Using that the map $C^{\alpha,0}([0,T];\mathbb{R}^d) \ni W \mapsto \mathbf{W} \in G\Omega_\alpha(\mathbb{R}^d)$ is measurable we have that the measure $\mathcal{L}^{\mathbf{W}}$ satisfies $\mathcal{L}^{\mathbf{W}} = \mathcal{L}^W \circ \mathbf{W}^{-1}$.

Secondly, for Gaussian rough paths $\mathbf{X}$ and $\mathbf{Y}$ that are α-Hölder continuous and some Reproducing Kernel Hilbert space rough path $\mathbf{h}$ that is β-Hölder continuous for $\alpha+\beta>1$, we have translation invariance of the homogenous rough path metric with respect to a Cameron Martin perturbation, namely

$$d_\alpha(\mathbf{Y}_{s,t},\mathbf{X}_{s,t}) = d_\alpha\Big(T^h(\mathbf{Y})_{s,t}, T^h(\mathbf{X})_{s,t}\Big).$$

Finally, we observe that the set $\{y \in C^{\alpha,0}([0,T];\mathbb{R}^d) : d_\alpha(\mathbf{W}(y),\mathbf{1}) < \varepsilon\}$ is symmetric around

0. Applying [KLL94, Theorem 2] gives

$$\begin{aligned}
\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{h},\varepsilon)\Big] =&\mathcal{L}^{W}\Big[\{x\in C^{\alpha,0}([0,T];\mathbb{R}^d): d_\alpha(\mathbf{W}(x),\mathbf{h})<\varepsilon\}\Big]\\
=&\mathcal{L}^{W}\Big[\{x\in C^{\alpha,0}([0,T];\mathbb{R}^d): d_\alpha(\mathbf{W}(x-h),\mathbf{1})<\varepsilon\}\Big]\\
=&\mathcal{L}^{W}\Big[\{y\in C^{\alpha,0}([0,T];\mathbb{R}^d): d_\alpha(\mathbf{W}(y),\mathbf{1})<\varepsilon\}+h\Big]\\
\geq&\exp\Big(\frac{-\|h\|_{\mathcal{H}}^2}{2}\Big)\mathcal{L}^{W}\Big[\{y\in C^{\alpha,0}([0,T];\mathbb{R}^d): d_\alpha(\mathbf{W}(y),\mathbf{1})<\varepsilon\}\Big]\\
\geq&\exp\Big(\frac{-\|h\|_{\mathcal{H}}^2}{2}\Big)\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1},\varepsilon)\Big].
\end{aligned}$$

□

Definition 15.1.3. *Let $\mathcal{L}^W$ be a Gaussian measure and let $\mathcal{L}^{\mathbf{W}}$ be the law of the lift to the Gaussian rough path. We define the Freidlin-Wentzell Large Deviations Rate Function by*

$$I^{\mathbf{W}}(\mathbf{X}):=\begin{cases}\frac{\|\pi_1(\mathbf{X})\|_{\mathcal{H}}^2}{2}, & \textit{if } \pi_1(\mathbf{X})\in\mathcal{H}\\ \infty, & \textit{otherwise.}\end{cases}$$
$$I^{\mathbf{W}}(\mathbf{X},\varepsilon):=\inf_{d_{\alpha;[0,T]}(\mathbf{X},\mathbf{Y})<\varepsilon} I(\mathbf{Y})$$

The following Corollary is similar to a result first proved in [LS01] for Gaussian measures.

Corollary 15.1.4. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 and let $\mathcal{L}^{\mathbf{W}}$ be the law of the lift to the Gaussian rough path. Then for $\mathbf{Y}\in\overline{S_2(\mathcal{H})}^{d_\alpha}$ and $a\in[0,1]$*

$$\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{Y},\varepsilon)\Big]\geq\exp\Big(I^{\mathbf{W}}(\mathbf{Y},a\varepsilon)\Big)\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha\big(\mathbf{1},(1-a)\varepsilon\big)\Big]$$

Proof. Using the fact that the lift of the RKHS is dense in the support of the Gaussian rough path, we know that there must exist at least one $h\in\mathcal{H}$ such that $d_\alpha(\mathbf{h},\mathbf{Y})<a\varepsilon$ for any choice of $a\in[0,1]$. Further, by nesting of sets

$$\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{Y},\varepsilon)\Big]\geq\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha\big(\mathbf{h},(1-a)\varepsilon\big)\Big].$$

Now apply Lemma 15.1.2 and take a minimum over all possible choices of $\mathbf{h}$. □

Before stating our main result of this section we recall a useful inequality stated in [CLL13].

Lemma 15.1.5 (Borell's rough path inequality)**.** *Let $\mathcal{L}^W$ be a Gaussian measure and let $\mathcal{L}^{\mathbf{W}}$ be the law of the lift to the Gaussian rough path. Let $\mathcal{K}\subset\mathcal{H}$ be the unit ball with respect to $\|\cdot\|_{\mathcal{H}}$ and denote $\mathbf{K}=\{\mathbf{h}=S_2[h]:h\in\mathcal{H}\}$.*

Let A be a Borel subset of $G\Omega_\alpha(\mathbb{R}^d)$, $\lambda>0$ and define

$$T[A,\delta_\lambda(\mathbf{K})]:=\Big\{T^h(\mathbf{X}):\mathbf{X}\in A,\mathbf{h}\in\delta_\lambda(\mathbf{K})\Big\}.$$

Then, denoting by $\mathcal{L}^{\mathbf{W}}_*$ the inner measure, we have

$$\mathcal{L}^{\mathbf{W}}_*\Big[T[A,\delta_\lambda(\mathbf{K})]\Big]\geq\Phi\Big(\lambda+\Phi^{-1}(\mathcal{L}^{\mathbf{W}}[A])\Big).$$

15.2 Gaussian Correlation Inequalities

A useful trick for proving small ball probabilities is Šidák's Lemma (see Lemma 15.2.1 below). It has been widely used to account for possible correlation between collections of normally distributed random variables. We are unable to use Šidák's Lemma in our context as we will be dealing both with the increments of a Gaussian measure and the increments of the iterated integrals of the path. Thus, we need to prove a new version of Šidák's Lemma that allows us to consider these terms in the higher levels of a Wiener-Itô chaos expansion.

Given an abstract Wiener space $(E, \mathcal{H}, \mathfrak{i})$, we consider the element $h \in \mathcal{H}$ as a random variable on the probability space $(E, \mathcal{B}(E), \mathcal{L})$ where $\mathcal{B}(E)$ is the cylindrical σ-algebra generated by the elements of E^*. When E is separable, $\mathcal{B}(E)$ is equal to the Borel σ-algebra.

Lemma 15.2.1 (Šidák's Lemma [Šid68, Kha67])**.** *Let $(E, \mathcal{H}, \mathfrak{i})$ be an abstract Wiener space with Gaussian measure $\mathcal{L}$ and let I be a countable index. Suppose $\forall j \in I$ that $h_j \in \mathcal{H}$ and $\varepsilon_j > 0$. Then for any $k \in I$*

$$\mathcal{L}\Big[\bigcap_{j\in I}\Big\{|h_j| < \varepsilon_j\Big\}\Big] \geq \mathcal{L}\Big[\Big\{|h_k| < \varepsilon_k\Big\}\Big]\mathcal{L}\Big[\bigcap_{j\in I\setminus\{k\}}\Big\{|h_j| < \varepsilon_j\Big\}\Big]. \tag{15.2.1}$$

Equivalently, for $h_k' \in \mathcal{H}$ such that $\|h_k\|_{\mathcal{H}} = \|h_k'\|_{\mathcal{H}}$ and $\forall j \in I\setminus\{k\}$, $\langle h_j, h_k'\rangle_{\mathcal{H}} = 0$ then

$$\mathcal{L}\Big[\bigcap_{j\in I}\Big\{|h_j| < \varepsilon_j\Big\}\Big] \geq \mathcal{L}\Big[\bigcap_{j\in I\setminus\{k\}}\Big\{|h_j| < \varepsilon_j\Big\} \cap \Big\{|h_k'| < \varepsilon_k\Big\}\Big]. \tag{15.2.2}$$

For an eloquent proof, see [Bog98, Theorem 4.10.3]. In particular, given a Gaussian process W, a countable collection of intervals $(s_j, t_j)_{j\in I}$ and bounds $(\varepsilon_j)_{j\in I}$, we have

$$\begin{aligned}\mathbb{P}\Big[\bigcap_{j\in I}\Big\{|W_{s_j,t_j}| < \varepsilon_j\Big\}\Big] \geq& \mathbb{P}\Big[\Big\{|W_{s_1,t_1}| < \varepsilon_1\Big\}\Big] \cdot \mathbb{P}\Big[\bigcap_{\substack{j\in I\\ j\neq 1}}\Big\{|W_{s_j,t_j}| < \varepsilon_j\Big\}\Big]\\ \geq& \prod_{j\in I}\mathbb{P}\Big[\Big\{|W_{s_j,t_j}| < \varepsilon_j\Big\}\Big].\end{aligned}$$

Thus the probability of a sequence of intervals of a Gaussian process sitting on slices is minimised when the Gaussian random variables are all independent.

This is an example of the now proved Gaussian corellation inequality (first proved in [Roy14]) which states

$$\mathbb{P}\Big[\bigcap_{j\in I}\Big\{|W_{s_j,t_j}| < \varepsilon_j\Big\}\Big] \geq \mathbb{P}\Big[\bigcap_{j\in I_1}\Big\{|W_{s_j,t_j}| < \varepsilon_j\Big\}\Big] \cdot \mathbb{P}\Big[\bigcap_{j\in I_2}\Big\{|W_{s_j,t_j}| < \varepsilon_j\Big\}\Big] \tag{15.2.3}$$

where $I_1 \cup I_2 = I$ and $I_1 \cap I_2 = \emptyset$.

Given a pair of abstract Wiener spaces $(E_1, \mathcal{H}_1, i_1)$ and $(E_2, \mathcal{H}_2, i_2)$, we can define a Gaussian measure on the Cartesian product $E_1 \times E_2$ which has Reproducing Kernel Hilbert space $\mathcal{H}_1 \times \mathcal{H}_2$ by taking the product measure $\mathcal{L}_1 \times \mathcal{L}_2$ over $(E_1 \times E_2, \mathcal{B}(E_1) \otimes \mathcal{B}(E_2))$.

We define the Tensor space $E_1 \otimes_\varepsilon E_2$ of E_1 and E_2 to be the closure of the algebraic tensor $E_1 \otimes E_2$ with respect to the injective tensor norm

$$\varepsilon(x) := \sup\Big\{|(f\otimes g)(x)| : f \in E_1^*, g \in E_2^*, \|f\|_{E_1^*} = \|g\|_{E_2^*} = 1\Big\}.$$

Let $f \in (E_1 \otimes_\varepsilon E_2)^*$. Then the map $E_1 \times E_2 \ni (x, y) \mapsto f(x \otimes y)$ is measurable and the pushforward of f with respect to the Gaussian measure is an element of the second Wiener Itô chaos. In the case where the tensor product is of two Hilbert spaces, there is no question over the choice of the norm for $\mathcal{H}_1 \otimes \mathcal{H}_2$.

A problem similar to this was first studied in [LQZ02]. We emphasise that our result is more general.

Lemma 15.2.2. *Let $(E_1, \mathcal{H}_1, \mathfrak{i}_1)$ and $(E_2, \mathcal{H}_2, \mathfrak{i}_2)$ be abstract Wiener spaces with Gaussian measures $\mathcal{L}_1$ and $\mathcal{L}_2$. Let $\mathcal{L}_1 \times \mathcal{L}_2$ be the product measure over the direct sum $E_1 \oplus E_2$. Let I_1, I_2, I_3 be countable indexes. Suppose that $\forall j \in I_1, h_{j,1} \in \mathcal{H}_1$ and $\varepsilon_{j,1} > 0$, $\forall j \in I_2, h_{j,2} \in \mathcal{H}_2$ and $\varepsilon_{j,2} > 0$, and $\forall j \in I_3, h_{j,3} \in \mathcal{H}_1 \otimes \mathcal{H}_2$ and $\varepsilon_{j,3} > 0$. Additionally, denote $\hat{\otimes} : E_1 \oplus E_2 \to E_1 \otimes_\varepsilon E_2$ by*

$\hat{\otimes}(x, y) = x \otimes y$. *Then*

$$(\mathcal{L}_1 \times \mathcal{L}_2)\Big[\bigcap_{j\in I_1}\Big\{|h_{j,1}| < \varepsilon_{j,1}\Big\}\bigcap_{j\in I_2}\Big\{|h_{j,2}| < \varepsilon_{j,2}\Big\}\bigcap_{j\in I_3}\Big\{|h_{j,3}(\hat{\otimes})| < \varepsilon_{j,3}\Big\}\Big]$$
$$\geq \prod_{j\in I_1}\mathcal{L}_1\Big[\Big\{|h_{j,1}| < \varepsilon_{j,1}\Big\}\Big]\cdot\prod_{j\in I_2}\mathcal{L}_2\Big[\Big\{|h_{j,2}| < \varepsilon_{j,2}\Big\}\Big]\cdot\prod_{j\in I_3}(\mathcal{L}_1 \times \mathcal{L}_2)\Big[\Big\{|h_{j,3}(\hat{\otimes})| < \varepsilon_{j,3}\Big\}\Big]$$

Proof. It should be clear that when $I_3 = \emptyset$, Lemma 15.2.2 comes immediately by applying Lemma 15.2.1. When $I_3 \neq \emptyset$, the bilinear forms $f_{j,3}(\hat{\otimes})$ are not bounded on $E_1 \times E_2$ and so we cannot immediately apply Lemma 15.2.1 (they are bounded on the space $E_1 \otimes_\varepsilon E_2$).

However, we do have that for $y \in E_2$ fixed, the functional $x \mapsto h(x \otimes y)$ is a linear functional and for $x \in E_1$ fixed, the functional $y \mapsto h(x \otimes y)$ is a linear functional (not necessarily bounded functionals). Thus, by the definition of the product measure, we have

$$\begin{aligned}
&(\mathcal{L}_1 \times \mathcal{L}_2)\Big[\bigcap_{j\in I_1}\Big\{|h_{j,1}| < \varepsilon_{j,1}\Big\}\bigcap_{j\in I_2}\Big\{|h_{j,2}| < \varepsilon_{j,2}\Big\}\bigcap_{j\in I_3}\Big\{|h_{j,3}(\hat{\otimes})| < \varepsilon_{j,3}\Big\}\Big]\\
&= \int_{E_2}\int_{E_1}\prod_{j\in I_1}\mathbb{1}_{\{|h_{j,1}|<\varepsilon_{j,1}\}}(x)\cdot\prod_{j\in I_2}\mathbb{1}_{\{|h_{j,2}|<\varepsilon_{j,2}\}}(y)\\
&\quad\cdot\prod_{j\in I_3}\mathbb{1}_{\{|h_{j,3}(\hat{\otimes})|<\varepsilon_{j,3}\}}(x,y)d\mathcal{L}_1(x)d\mathcal{L}_2(y)\\
&\geq \int_{E_2}\prod_{j\in I_2}\mathbb{1}_{\{|h_{j,2}|<\varepsilon_{j,2}\}}(y)\cdot\int_{E_1}\prod_{j\in I_1}\mathbb{1}_{\{|h'_{j,1}|<\varepsilon_{j,1}\}}(x)\\
&\quad\cdot\prod_{j\in I_3}\mathbb{1}_{\{|h'_{j,3}(\hat{\otimes})|<\varepsilon_{j,3}\}}(x,y)d\mathcal{L}_1(x)d\mathcal{L}_2(y)
\end{aligned}$$

where for each $j \in I_1$ $\|h'_{j,1}\|_{\mathcal{H}} = \|h_{j,1}\|_{\mathcal{H}}$, for each $j \in I_3$ and $y \in E_2$ fixed $\|h'_{j,3}(\cdot \otimes y)\|_{\mathcal{H}} = \|h_{j,3}(\cdot \otimes y)\|_{\mathcal{H}}$, and the vectors $\{h'_{j,1}\}_{j\in I_1} \cup \{h'_{j,3}(\cdot \otimes y)\}_{j\in I_3}$ are orthonormal in $\mathcal{H}$. This comes from applying Equation (15.2.2) from Lemma 15.2.1.

Similarly, swapping the order of integration and repeating yields

$$\geq \int_{E_1}\prod_{j\in I_1}\mathbb{1}_{\{|h'_{j,1}|<\varepsilon_{j,1}\}}(x)\cdot\int_{E_2}\prod_{j\in I_2}\mathbb{1}_{\{|h'_{j,2}|<\varepsilon_{j,2}\}}(y)\cdot\prod_{j\in I_3}\mathbb{1}_{\{|h''_{j,3}(\hat{\otimes})|<\varepsilon_{j,3}\}}(x,y)d\mathcal{L}_2(y)d\mathcal{L}_1(x)$$
$$\geq \prod_{j\in I_1}\mathcal{L}_1\Big[\Big\{|h'_{j,1}| < \varepsilon_{j,1}\Big\}\Big]\cdot\prod_{j\in I_2}\mathcal{L}_2\Big[\Big\{|h'_{j,2}| < \varepsilon_{j,2}\Big\}\Big]\cdot\prod_{j\in I_3}(\mathcal{L}_1 \times \mathcal{L}_2)\Big[\Big\{|h''_{j,3}(\hat{\otimes})| < \varepsilon_{j,3}\Big\}\Big]$$

where for each $j \in I_2$ $\|h'_{j,2}\|_{\mathcal{H}} = \|h_{j,2}\|_{\mathcal{H}}$, for each $j \in I_3$ and $x \in E_1$ fixed $\|h'_{j,3}(x \otimes \cdot)\|_{\mathcal{H}} = \|h''_{j,3}(x \otimes \cdot)\|_{\mathcal{H}}$, and the vectors $\{h'_{j,2}\}_{j\in I_2} \cup \{h''_{j,3}(x \otimes \cdot)\}_{j\in I_3}$ are orthonormal in $\mathcal{H}$. $\square$

In fact, rather than dividing this intersection of sets into a product of probabilities completely (as will be necessary later in this chapter), we could have used Equation (15.2.3) to divide the intersection into the product of any number of two intersections. We do not state this to avoid writing already challenging notation and because there is no need for such a result in Chapter 16.

Proposition 15.2.3 (Šidák's Lemma for Higher order Wiener-Itô chaos elements.)**.** *Let m be a positive integer. Let $(E_1, \mathcal{H}_1, \mathrm{i}_1)$, ..., $(E_m, \mathcal{H}_m, \mathrm{i}_m)$ be m Abstract Wiener spaces with Gaussian measures $\mathcal{L}_1$, ..., $\mathcal{L}_m$. Let $\mathcal{L}_1 \times ... \times \mathcal{L}_m$ be the product measure over the direct sum $E_1 \oplus ... \oplus E_m$. Let I_1, I_2, ..., I_m be m countable indexes. Suppose that for $l \in \{1, ..., m\}$, $\forall j \in I_l$*

$$h_{j,l} \in \bigcup_{\substack{k_1,...,k_l\\ k_1\neq...\neq k_l}}\mathcal{H}_{k_1}\otimes ...\otimes\mathcal{H}_{k_l}, \quad \varepsilon_{j,l} > 0.$$

Next, suppose

$$\hat{\otimes}_l : E_{k_1} \oplus ... \oplus E_{k_l} \to E_{k_1} \otimes_\varepsilon ... \otimes_\varepsilon E_{k_l}, \quad \hat{\otimes}(x_{k_1}, ..., x_{k_l}) := x_{k_1} \otimes ... \otimes x_{k_l}$$

Then

$$\Big(\mathcal{L}_1 \times ... \times \mathcal{L}_m\Big)\left[\bigcap_{l=1}^{m} \bigcap_{j \in I_l} \Big\{|h_{j,l}(\hat{\otimes}_l)| < \varepsilon_{j,l}\Big\}\right] \geq \prod_{l=1}^{m} \prod_{j \in I_l} \Big(\mathcal{L}_1 \times ... \times \mathcal{L}_m\Big)\left[\Big\{|h_{j,l}(\hat{\otimes}_l)| < \varepsilon_{j,l}\Big\}\right]$$

Proof. Repetitive applications of the methods of Lemma 15.2.2 and Equation (15.2.2). □

Chapter 16

Small Ball Probabilities for Enhanced Gaussian Processes

The aim of this Chapter is to prove the following two Theorems.

Theorem 16.0.1. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 for some $\varrho \in [1, 3/2)$ and let $\mathbf{W}$ be the lifted Gaussian rough path. Then for $\frac{1}{3} < \alpha < \frac{1}{2\varrho}$ we have*

$$\mathfrak{B}(\varepsilon) = -\log\Big(\mathbb{P}\Big[\|\mathbf{W}\|_\alpha < \varepsilon\Big]\Big) \lesssim \varepsilon^{\frac{-1}{\frac{1}{2\varrho}-\alpha}}. \tag{16.0.1}$$

Assumption 16.0.2. *Let $\mathcal{L}^W$ be a Gaussian measure.*

1. *Suppose that either W satisfies*
   - $\mathbb{E}\Big[|W_{s,t}|^2\Big] = \varsigma\Big(|t-s|\Big)$,
   - $\varsigma\Big(|t-s|\Big) \geq |t-s|^{\frac{1}{\varrho}}$,
   - $\exists c > 0$ *such that* $\varsigma(2t) \leq c\varsigma(t)$ *and*
   - $\Big|\frac{d^4\varsigma}{dt^4}(t)\Big| \leq \frac{M}{t^{4-\frac{1}{\varrho}}}$.
2. *Alternatively suppose that*
$$\mathcal{R}_{s,t} = \mathbb{E}\Big[|W_{s,t}|^2\Big] = |t-s|^{\frac{1}{\varrho}}.$$

Theorem 16.0.3. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 for some $\varrho \in [1, 3/2)$. Additionally, suppose that Assumption 16.0.2 holds. Then*

$$\mathfrak{B}(\varepsilon) = -\log\Big(\mathbb{P}\Big[\|\mathbf{W}\|_\alpha < \varepsilon\Big]\Big) \gtrsim \varepsilon^{\frac{-1}{\frac{1}{2\varrho}-\alpha}}$$

These results can be found in the preprint [Sal20, Section 4].

16.1 Norm Discretisation

Firstly, we address a method for discretising the rough path Hölder norm. To the best of the authors knowledge, this result has not previously been stated in the framework of rough paths. The proof is an adaption of the tools used in [KLS95, Theorem 2.2].

Lemma 16.1.1 (Discretisation of rough path norms)**.** *Let* $\mathbf{W} \in WG\Omega_\alpha(\mathbb{R}^d)$ *be a rough path. Then we have*

$$\|\mathbf{W}\|_\alpha \leq \max\left(2\sum_{l=1}^{\infty}\sup_{i=1,\ldots,2^l}\frac{\|\mathbf{W}_{(i-1)T2^{-l},iT2^{-l}}\|_{cc}}{\varepsilon^\alpha},\right.$$
$$\left. 3\sup_{j\in\mathbb{N}_0}\sup_{i=0,\ldots\left\lfloor\frac{2^j(T-\varepsilon)}{\varepsilon}\right\rfloor}\sum_{l=j+1}^{\infty}\sup_{m=1,\ldots,2^{l-j}}\frac{\|\mathbf{W}_{(m-1)2^{-l}+i\varepsilon 2^{-j},m2^{-l}+i\varepsilon 2^{-j}}\|_{cc}}{\varepsilon^\alpha 2^{-\alpha(j+1)}}\right).$$

Proof. Let $0 < \varepsilon < T$.

$$\|\mathbf{W}\|_\alpha \leq \sup_{\substack{s,t\in[0,T]\\|t-s|\geq\varepsilon}}\frac{\|\mathbf{W}_{s,t}\|_{cc}}{|t-s|^\alpha}\bigvee\sup_{\substack{s,t\in[0,T]\\|t-s|<\varepsilon}}\frac{\|\mathbf{W}_{s,t}\|_{cc}}{|t-s|^\alpha}\leq\sup_{s\in[0,T]}\frac{2\|\mathbf{W}_{0,s}\|_{cc}}{\varepsilon^\alpha}\bigvee\sup_{\substack{0\leq s\leq T\\0\leq t\leq\varepsilon\\|s+t|<T}}\frac{\|\mathbf{W}_{s,s+t}\|_{cc}}{|t|^\alpha}\tag{16.1.1}$$

Firstly, writing $s \in [0,T]$ as a sum of dyadics and exploiting the sub-additivity of the Carnot-Carathéodory norm, we get

$$\|\mathbf{W}_{0,s}\|_{cc} \leq \sum_{l=1}^{\infty}\sup_{i=1,\ldots,2^l}\|\mathbf{W}_{(i-1)T2^{-l},iT2^{-l}}\|_{cc}.\tag{16.1.2}$$

Hence

$$\sup_{s\in[0,T]}\frac{2\|\mathbf{W}_{0,s}\|_{cc}}{\varepsilon^\alpha}\leq\sum_{l=1}^{\infty}\sup_{i=1,\ldots,2^l}\frac{2\|\mathbf{W}_{(i-1)T2^{-l},iT2^{-l}}\|_{cc}}{\varepsilon^\alpha}.\tag{16.1.3}$$

Secondly,

$$\begin{aligned}\sup_{\substack{0\leq s\leq T\\0\leq t\leq\varepsilon}}\frac{\|\mathbf{W}_{s,s+t}\|_{cc}}{|t|^\alpha} &\leq \sup_{s\in[0,T]}\max_{j\in\mathbb{N}_0}\sup_{\varepsilon 2^{-j-1}\leq t<\varepsilon 2^{-j}}\frac{\|\mathbf{W}_{s,s+t}\|_{cc}}{|t|^\alpha}\\ &\leq \max_{j\in\mathbb{N}_0}\sup_{s\in[0,T]}\sup_{0<t<\varepsilon 2^{-j}}\frac{\|\mathbf{W}_{s,s+t}\|_{cc}}{|\varepsilon|^\alpha\cdot 2^{-\alpha(j+1)}}\\ &\leq \max_{j\in\mathbb{N}_0}\max_{i=0,\ldots,\left\lfloor\frac{2^j(T-\varepsilon)}{\varepsilon}\right\rfloor}\sup_{0<t<\varepsilon 2^{-j}}\frac{3\|\mathbf{W}_{i\varepsilon 2^{-j},i\varepsilon 2^{-j}+t}\|_{cc}}{|\varepsilon|^\alpha\cdot 2^{-\alpha(j+1)}}.\end{aligned}$$

Then, as with Equation (16.1.2) we have that for $t \in (0, \varepsilon 2^{-j})$,

$$\|\mathbf{W}_{i\varepsilon 2^{-j},i\varepsilon 2^{-j}+t}\|_{cc}\leq\sum_{l=j+1}^{\infty}\sup_{m=1,\ldots,2^{l-j}}\|\mathbf{W}_{(m-1)2^{-l}+i\varepsilon 2^{-j},m2^{-l}+i\varepsilon 2^{-j}}\|_{cc}.$$

Hence

$$\sup_{\substack{0\leq s\leq T\\0\leq t\leq\varepsilon\\|s+t|<T}}\frac{\|\mathbf{W}_{s,s+t}\|_{cc}}{|t|^\alpha}\leq\sup_{j\in\mathbb{N}_0}\sup_{i=0,\ldots\left\lfloor\frac{2^j(T-\varepsilon)}{\varepsilon}\right\rfloor}\sum_{l=j+1}^{\infty}\sup_{m=1,\ldots,2^{l-j}}\frac{3\|\mathbf{W}_{(m-1)2^{-l}+i\varepsilon 2^{-j},m2^{-l}+i\varepsilon 2^{-j}}\|_{cc}}{\varepsilon^\alpha 2^{-\alpha(j+1)}}.\tag{16.1.4}$$

Combining Equation (16.1.1) with Equation (16.1.3) and Equation (16.1.4) yields the result. □

16.2 Proofs of Main Results

Proof of Theorem 16.0.1. Let n_0 be a positive integer such that $\varepsilon^{-1} \leq 2^{n_0} \leq 2\varepsilon^{-1}$ and denote $\beta = \frac{1}{2\varrho} - \alpha$ for brevity. Define

$$\begin{aligned} \varepsilon_l^{(1)} :=& \Big(\frac{3}{2}\Big)^{\frac{-|l-n_0|}{2\varrho}} \varepsilon^{\frac{1}{2\varrho}} \cdot \frac{(1-2^{-\beta/2})}{4}, \\ \varepsilon_{j,l}^{(2)} :=& \frac{\varepsilon^\beta 2^{\frac{-l}{2\varrho}}}{3} \cdot \frac{2^{\beta(l+j)/2}(1-2^{-\beta/2})}{2^{-\alpha(j+1)}}. \end{aligned} \tag{16.2.1}$$

Observe that these satisfy the properties

$$\sum_{l=1}^\infty \varepsilon_l^{(1)} \leq \frac{\varepsilon^{\frac{1}{2\varrho}}}{2} \quad \text{and} \quad \sum_{l=j+1}^\infty \varepsilon_{j,l}^{(2)} \leq \frac{\varepsilon^\beta}{3}.$$

Therefore, using Lemma 16.1.1 gives the lower bound

$$\begin{aligned} \mathbb{P}\Big[\|\mathbf{W}\|_\alpha \leq \varepsilon^\beta\Big] \geq \mathbb{P}\Bigg[& \sup_{i=1,\dots,2^l} \|\mathbf{W}_{(i-1)T2^{-l},iT2^{-l}}\|_{cc} \leq \varepsilon_l^{(1)} \quad \forall l \in \mathbb{N}, \\ & \sup_{i=0,\dots,\left\lfloor \frac{2^j(T-\varepsilon)}{\varepsilon}\right\rfloor} \sup_{m=1,\dots,2^{l-j}} \frac{\|\mathbf{W}_{(m-1)2^{-l}\varepsilon+i2^{-j}\varepsilon,m2^{-l}\varepsilon+i2^{-l}\varepsilon}\|_{cc}}{\varepsilon^\alpha 2^{-\alpha(j+1)}} \leq \varepsilon_{j,l}^{(2)} \quad \forall l \geq j+1, j,l \in \mathbb{N}_0\Bigg]. \end{aligned} \tag{16.2.2}$$

Next, using the equivalence of the Homogeneous norm from Equation C.1.1, we have that there exists a constant dependent only on d such that

$$\|\mathbf{W}_{s,t}\|_{cc} \leq c(d) \cdot \sup_{A \in \mathcal{A}_2} \Big|\langle \log_\boxtimes(\mathbf{W}_{s,t}), e_A\rangle\Big|^{1/|A|}.$$

Using the Philip-Hall Lie basis for the Lie Algebra $\log_\boxtimes(G^2(\mathbb{R}^d))$, we get the representation

$$\|\mathbf{W}_{s,t}\|_{cc} \leq c(d) \sup_{p=1,\dots,d} \Big|\Big\langle \mathbf{W}_{s,t}, e_{(p)}\Big\rangle\Big| \bigvee \sup_{\substack{p,q=1,\dots,d\\ p\neq q}} \Big|\Big\langle \mathbf{W}_{s,t}, e_{(p,q)}\Big\rangle\Big|^{1/2}.$$

Applying Proposition 15.2.3 to this yields

$$\begin{aligned} &\mathbb{P}\Big[\|\mathbf{W}\|_\alpha \leq \varepsilon^\beta\Big] \\ &\geq \Bigg\{ \prod_{l=1}^\infty \prod_{i=1}^{2^l} \Bigg(\prod_{p=1}^d \mathbb{P}\Big[\Big|\Big\langle \mathbf{W}_{(i-1)T2^{-l},iT2^{-l}}, e_{(p)}\Big\rangle\Big| \leq \tfrac{\varepsilon_l^{(1)}}{c(d)}\Big] \\ &\quad \cdot \prod_{\substack{p,q=1\\ p\neq q}}^d \mathbb{P}\Big[\Big|\Big\langle \mathbf{W}_{(i-1)T2^{-l},iT2^{-l}}, e_{(p,q)}\Big\rangle\Big| \leq \Big(\tfrac{\varepsilon_l^{(1)}}{c(d)}\Big)^2\Big] \\ &\quad \times \Bigg\{ \prod_{j=0}^\infty \prod_{l=j+1}^\infty \prod_{i=0}^{\lfloor 2^j(T-\varepsilon)/\varepsilon \rfloor} \prod_{m=1}^{2^{l-j}} \Bigg(\prod_{p=1}^d \mathbb{P}\Big[\Big|\Big\langle \mathbf{W}_{\varepsilon(m-1)2^{-l}+\varepsilon i2^{-j},\varepsilon m2^{-l}+\varepsilon i2^{-j}}, e_{(p)}\Big\rangle\Big| \leq \tfrac{\varepsilon_{j,l}^{(2)} \cdot \varepsilon^\alpha 2^{-\alpha(j+1)}}{c(d)}\Big] \\ &\quad \cdot \prod_{\substack{p,q=1\\ p\neq q}}^d \mathbb{P}\Big[\Big|\Big\langle \mathbf{W}_{\varepsilon(m-1)2^{-l}+\varepsilon i2^{-j},\varepsilon m2^{-l}+\varepsilon i2^{-j}}, e_{(p,q)}\Big\rangle\Big| \leq \Big(\tfrac{\varepsilon_{j,l}^{(2)} \cdot \varepsilon^\alpha 2^{-\alpha(j+1)}}{c(d)}\Big)^2\Big] \end{aligned} \tag{16.2.3}$$

For the terms associated to words of length 1, the computation of this probability under

Assumption C.2.12 is simply

$$\mathbb{P}\Big[|W_{s,t}| \leq \varepsilon\Big] = \operatorname{erf}\left(\tfrac{\varepsilon}{\sqrt{2}\mathbb{E}[|W_{s,t}|^2]^{1/2}}\right) \geq \operatorname{erf}\left(\tfrac{\varepsilon}{\sqrt{2}M|t-s|^{1/2\varrho}}\right). \tag{16.2.4}$$

For longer words, we only attain the lower bound.

$$\begin{aligned}\mathbb{P}\Big[\Big|\int_s^t W_{s,r}^{(p)}dW_r^{(q)}\Big| < \varepsilon\Big] =&\mathbb{E}\left[\mathbb{P}\left[\Big|\int_s^t W_{s,r}^{(p)}dW_r^{(q)}\Big| < \varepsilon\middle|\sigma\Big(W^{(p)}\Big)\right]\right]\\ =&\mathbb{E}\Big[\operatorname{erf}\left(\tfrac{\varepsilon}{\sqrt{2}\|W_{s,\cdot}\mathbb{1}_{s,t}\|_{\mathcal{H}}}\right)\Big]\\ \geq& \operatorname{erf}\left(\tfrac{\varepsilon}{\sqrt{2}\mathbb{E}[\|W_{s,\cdot}\mathbb{1}_{s,t}\|_{\mathcal{H}}^2]^{1/2}}\right) \geq \operatorname{erf}\left(\tfrac{\varepsilon}{\sqrt{2}M|t-s|^{2/(2\varrho)}}\right).\end{aligned} \tag{16.2.5}$$

We also use the lower bounds

$$\operatorname{erf}\left(\tfrac{t}{\sqrt{2}}\right) \geq \tfrac{t}{2} \qquad \text{for } t\in[0,1], \tag{16.2.6}$$

$$\operatorname{erf}\left(\tfrac{st}{\sqrt{2}}\right) \geq \exp\left(\frac{-\exp\left(\frac{-(st)^2}{2}\right)}{1-\exp\left(\frac{-s^2}{2}\right)}\right) \qquad \text{for } s>0, t\in[1,\infty). \tag{16.2.7}$$

We now consider the terms from Equation 16.2.3 with the product over (j,l,i,m). By Assumption C.2.12, the expression (16.2.1) and Equation 16.2.4 we have

$$\mathbb{P}\left[\Big|\Big\langle \mathbf{W}_{\varepsilon(m-1)2^{-l}+\varepsilon i2^{-j},\varepsilon m2^{-l}+\varepsilon i2^{-j}}, e_{(p)}\Big\rangle\Big| \leq \frac{\varepsilon_{j,l}^{(2)}\cdot\varepsilon^\alpha 2^{-\alpha(j+1)}}{c(d)}\right] \geq \operatorname{erf}\left(\tfrac{(1-2^{-\beta/2})}{3Mc(d)}\cdot 2^{\beta(l+j)/2}\right).$$

By similarly applying Equation 16.2.5

$$\mathbb{P}\left[\Big|\Big\langle \mathbf{W}_{\varepsilon(m-1)2^{-l}+\varepsilon i2^{-j},\varepsilon m2^{-l}+\varepsilon i2^{-j}}, e_{(p,q)}\Big\rangle\Big| \leq \left(\frac{\varepsilon_{j,l}^{(2)}\cdot\varepsilon^\alpha 2^{-\alpha(j+1)}}{c(d)}\right)^2\right] \geq \operatorname{erf}\left(\left(\tfrac{(1-2^{-\beta/2})}{3Mc(d)}\right)^2\cdot 2^{2\beta(l+j)/2}\right).$$

Next, we denote $s=\frac{(1-2^{-\beta/2})}{3Mc(d)}$, apply the lower bound (16.2.7) and multiply all the terms together correctly to obtain

$$\begin{aligned}&\prod_{j=0}^\infty\prod_{l=j+1}^\infty \exp\left(\frac{-T2^l}{\varepsilon}\cdot\left[\frac{d}{1-e^{-s^2/2}}\exp\left(\tfrac{-s^2}{2}2^{\beta(l+j)}\right)+\frac{d(d-1)}{2(1-e^{-s^4/2})}\exp\left(\tfrac{-s^4}{2}2^{2\beta(l+j)}\right)\right]\right)\\ &\qquad\geq \exp\left(-\frac{c_1(d,T,M,\beta)}{\varepsilon}\right).\end{aligned} \tag{16.2.8}$$

Secondly, we consider the terms from Equation 16.2.3 with the product over (l,i) and restrict ourselves to the case where $l>n_0$. By applying the definition of n_0, $\varepsilon_l^{(1)}$ and using Assumption C.2.12

$$\mathbb{P}\left[\Big|\Big\langle \mathbf{W}_{(i-1)2^{-l},i2^{-l}}, e_{(p)}\Big\rangle\Big| \leq \tfrac{\varepsilon_l^{(1)}}{c(d)}\right] \geq \operatorname{erf}\left(\left(\tfrac{4}{3}\right)^{\frac{l-n_0}{2\varrho}}\cdot\tfrac{(1-2^{-\beta/2})}{4Mc(d)\sqrt{2}}\right).$$

Similarly, by using Equations (16.2.5),

$$\mathbb{P}\left[\Big|\Big\langle \mathbf{W}_{(i-1)2^{-l},i2^{-l}}, e_{(p,q)}\Big\rangle\Big| \leq \left(\tfrac{\varepsilon_l^{(1)}}{c(d)}\right)^2\right] \geq \operatorname{erf}\left(\left(\tfrac{4}{3}\right)^{\frac{2(l-n_0)}{2\varrho}}\cdot\tfrac{1}{\sqrt{2}}\cdot\left(\tfrac{(1-2^{-\beta/2})}{4Mc(d)}\right)^2\right),$$

Now applying Equation (16.2.7) and multiplying all the terms together gives

$$\prod_{l=n_0+1}^{\infty} \exp\Bigg(-2^l \Bigg[\frac{d}{1-e^{-s^2/2}} \exp\Big(-\tfrac{s^2}{2}\Big(\tfrac{4}{3}\Big)^{\frac{l-n_0}{2\varrho}} \Big) + \frac{d(d-1)}{2(1-e^{-s^4/2})} \exp\Big(-\tfrac{s^4}{2}\Big(\tfrac{4}{3}\Big)^{\frac{2(l-n_0)}{2\varrho}} \Big) \Bigg]\Bigg)$$
$$\geq \exp\Big(-2^{n_0} c_2(d,T,M,\beta)\Big) \geq \exp\Big(-\tfrac{2c_2(d,T,M,\beta)}{\varepsilon}\Big) \tag{16.2.9}$$

where $s = \Big(\frac{(1-2^{-\beta/2})}{4Mc(d)}\Big)$.

Finally, we come to the terms from Equation 16.2.3 with the product over (l,i) where we consider the remaining terms for $l = 0, ..., n_0$. Using the definition of ε and Assumption C.2.12

$$\mathbb{P}\bigg[\Big| \Big\langle \mathbf{W}_{(i-1)2^{-l}, i2^{-l}}, e_{(p)} \Big\rangle \Big| \leq \tfrac{\varepsilon_l^{(1)}}{c(d)} \bigg] \geq \operatorname{erf}\bigg(\Big(\tfrac{1}{3}\Big)^{\frac{n_0-l}{2\varrho}} \cdot \tfrac{1}{\sqrt{2}} \cdot \tfrac{1-2^{-\beta/2}}{4Mc(d)} \bigg).$$

Similarly, by using Equations (16.2.5),

$$\mathbb{P}\bigg[\Big| \Big\langle \mathbf{W}_{(i-1)2^{-l}, i2^{-l}}, e_{(p,q)} \Big\rangle \Big| \leq \Big(\tfrac{\varepsilon_l^{(1)}}{c(d)}\Big)^2 \bigg] \geq \operatorname{erf}\bigg(\Big(\tfrac{1}{3}\Big)^{\frac{2(n_0-l)}{2\varrho}} \cdot \tfrac{1}{\sqrt{2}} \cdot \Big(\tfrac{1-2^{-\beta/2}}{4Mc(d)}\Big)^2 \bigg),$$

For these terms, we use the lower bound (16.2.6) and multiply all the terms together to get

$$\prod_{l=0}^{n_0} \Bigg(\bigg[\tfrac{1}{2} \cdot \tfrac{(1-2^{-\beta/2})}{4Mc(d)} \cdot \Big(\tfrac{1}{3}\Big)^{\frac{(n_0-l)}{2\varrho}} \bigg]^{dT2^l} \cdot \bigg[\tfrac{1}{2} \cdot \Big(\tfrac{(1-2^{-\beta/2})}{4Mc(d)}\Big)^2 \cdot \Big(\tfrac{1}{3}\Big)^{\frac{2(n_0-l)}{2\varrho}} \bigg]^{\frac{d(d-1)T2^l}{2}}$$
$$\geq \exp\Bigg(-2^{n_0} \sum_{l=1}^{n_0} \bigg[dT\Big(2^{-(n_0-l)} \log\Big(2 \cdot \Big(\tfrac{4Md(c)}{1-2^{-\beta/2}}\Big)\Big) + (n_0-l)\log\Big(3^{\frac{1}{2\varrho}}\Big)\Big)$$
$$+ \tfrac{d(d-1)}{2}\Big(2^{-(n_0-l)} \log\Big(2 \cdot \Big(\tfrac{4Md(c)}{1-2^{-\beta/2}}\Big)^2\Big) + (n_0-l)\log\Big(3^{\frac{2}{2\varrho}}\Big)\Big)$$
$$\geq \exp\Big(-2^{n_0} c_3(d,T,M,\beta)\Big) \geq \exp\Big(-\tfrac{2c_3(d,T,M,\beta)}{\varepsilon}\Big). \tag{16.2.10}$$

Combining Equations (16.2.8), (16.2.9) and (16.2.10) gives that

$$(16.2.3) \geq \exp\Big(-\tfrac{(c_1+c_2+c_3)}{\varepsilon}\Big) \quad \Rightarrow \quad -\log\Big(\mathbb{P}\Big[\|\mathbf{W}\|_\alpha \leq \varepsilon^\beta\Big]\Big) \lesssim \varepsilon^{-1}.$$

□

Proof of Theorem 16.0.3. The lower bound is a simple exercise which we only explain briefly. We observe that

$$\mathbb{P}\Big[\|\mathbf{W}\|_\alpha < \varepsilon\Big] \leq \mathbb{P}\Big[\|W\|_\alpha < \varepsilon'\Big]$$

due to Equation (C.1.1). Then we use either [Sto96, Theorem 1.2 or Theorem 1.4] extended to the multi-dimensional setting. □

Chapter 17

Metric Entropy and Applications

This Chapter aims to demonstrate some of the applications of the Small Ball Probability results proved in Chapter 16. For the most part, the results of this Chapter are adaptions of previously known results to the rough path framework and demonstrate that the compactness properties of Gaussian processes are retained by taking the signature of the path. This is not immediately obvious and somewhat remarkable given that the map from a path to its signature is not even continuous. These results can be found in the preprint [Sal20, Section 5].

17.1 Metric Entropy of Cameron Martin Balls

This problem was first studied in [KL93a] for Gaussian measures. While the law of a Gaussian rough path has many of the properties that Gaussian measures are known for, it is not itself a Gaussian so this result is not immediate.

Definition 17.1.1. *Let (E, d) be a metric space and let K be a compact subset of E. We define the d-metric Entropy of K to be $\mathfrak{H}(\varepsilon, K) := \log(\mathfrak{N}(\varepsilon, K))$ where*

$$\mathfrak{N}(\varepsilon, K) := \min\left\{ n \geq 1 : \exists e_1, ..., e_n \in E, \bigcup_{j=1}^{n} \mathbb{B}(e_j, \varepsilon) \supseteq K \right\}$$

and $\mathbb{B}(e_i, \varepsilon) := \{e \in E : d(e, e_i) < \varepsilon\}$.

Given a Gaussian measure $\mathcal{L}^W$ with Reproducing Kernel Hilbert space $\mathcal{H}$ and unit ball $\mathcal{K}$, let us consider the set of rough paths

$$\mathbf{K} := \Big\{ \mathbf{h} = S_2[h] : h \in \mathcal{K} \Big\} \subset G\Omega_\alpha(\mathbb{R}^d). \tag{17.1.1}$$

We can easily show that this set is Equicontinuous as a path on $G^2(\mathbb{R}^d)$ so by the Arzelà–Ascoli theorem, see for example [FV10b, Theorem 1.4], it must be compact in the metric space $G\Omega_\alpha(\mathbb{R}^d)$. Hence $\mathfrak{N}_{d_\alpha}(\varepsilon, \mathbf{K})$ is finite.

Theorem 17.1.2. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 with Reproducing Kernel Hilbert space $\mathcal{H}$ and unit ball $\mathcal{K}$.*

Then the Metric Entropy of the set $\mathbf{K}$ with respect to the Hölder metric satisfies

$$\mathfrak{H}_{d_\alpha}(\varepsilon, \mathbf{K}) \lesssim \varepsilon^{\frac{-1}{\frac{1}{2}+\frac{1}{2\varrho}-\alpha}}.$$

Further, suppose that $\mathcal{L}^W$ satisfies Assumption 16.0.2. Then the Metric Entropy of the set $\mathbf{K}$ with respect to the Hölder metric additionally satisfies

$$\mathfrak{H}_{d_\alpha}(\varepsilon, \mathbf{K}) \gtrsim \varepsilon^{\frac{-1}{\frac{1}{2}+\frac{1}{2\varrho}-\alpha}}.$$

17.1.1 Auxilliary compactness results

In order to prove this, we first prove the following auxiliary Proposition:

Proposition 17.1.3. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 with Reproducing Kernel Hilbert space $\mathcal{H}$ and unit ball $\mathcal{K}$. Then for any $\eta, \varepsilon > 0$,*

$$\mathfrak{H}_{d_\alpha}\Big(2\varepsilon, \delta_\eta(\mathbf{K})\Big) \leq \tfrac{\eta^2}{2} - \log\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon)\Big]\Big), \tag{17.1.2}$$

and

$$\mathfrak{H}_{d_\alpha}\Big(\varepsilon, \delta_\eta(\mathbf{K})\Big) \geq \log\Bigg(\Phi\Big(\eta + \Phi^{-1}\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon)\Big]\Big)\Big)\Bigg) - \log\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, 2\varepsilon)\Big]\Big). \tag{17.1.3}$$

Proof. Firstly, for some $\varepsilon > 0$ consider the quantity

$$\mathfrak{M}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K})) = \max\Big\{n \geq 1 : \exists \mathbf{h}_1, ..., \mathbf{h}_n \in \delta_\eta(\mathbf{K}), d_\alpha(\mathbf{h}_i, \mathbf{h}_j) \geq 2\varepsilon \quad \forall i \neq j\Big\},$$

and a set $\mathfrak{F}$ such that $|\mathfrak{F}| = \mathfrak{M}_{\delta_\alpha}(\varepsilon, \delta_\eta(\mathbf{K}))$ and for any two distinct $\mathbf{h}_1, \mathbf{h}_2 \in \mathfrak{F}$ that $d_\alpha(\mathbf{h}_1, \mathbf{h}_2) \geq 2\varepsilon$. Similarly, there must exist a set $\mathfrak{G}$ such that $|\mathfrak{G}| = \mathfrak{N}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K}))$ and

$$\delta_\eta(\mathbf{K}) \subseteq \bigcup_{\mathbf{h} \in \mathfrak{G}} \mathbb{B}_\alpha(\mathbf{h}, \varepsilon).$$

Similarly, since $\mathfrak{F}$ is a maximal set, we also have

$$\delta_\eta(\mathbf{K}) \subseteq \bigcup_{\mathbf{h} \in \mathfrak{F}} \mathbb{B}_\alpha(\mathbf{h}, 2\varepsilon),$$

it is therefore natural that

$$\mathfrak{H}\Big(2\varepsilon, \delta_\eta(\mathbf{K})\Big) \leq \log\Big(\mathfrak{M}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K}))\Big) \quad \text{and} \quad \mathfrak{M}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K})) \min_{\mathbf{h} \in \mathfrak{F}} \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{h}, \varepsilon)\Big] \leq 1.$$

By taking Logarithms and applying Lemma 15.1.2 we get

$$\log\Big(\mathfrak{M}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K}))\Big) - \tfrac{\eta^2}{2} + \log\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon)\Big]\Big) \leq 0,$$

which implies (17.1.2).

Secondly, from the definition of $\mathfrak{F}$ we get that

$$T\Big(\mathbb{B}_\alpha(\mathbf{1}, \varepsilon), \delta_\eta(\mathbf{K})\Big) \subseteq \bigcup_{\mathbf{h} \in \mathfrak{G}} \mathbb{B}_\alpha(\mathbf{h}, 2\varepsilon),$$

where the set on the LHS satisfies

$$T\Big(\mathbb{B}_\alpha(\mathbf{1}, \varepsilon), \delta_\eta(\mathbf{K})\Big) = \Big\{T^{\eta h}(\mathbf{X}) : \mathbf{X} \in \mathbb{B}_\alpha(\mathbf{1}, \varepsilon), h \in \mathcal{K}\Big\}.$$

Additionally, we have for any choice of $\mathbf{h} = S_2[h]$, $h \in \mathcal{H}$, that

$$\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon)\Big] \geq \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{h}, \varepsilon)\Big].$$

Hence applying Lemma 15.1.5 gives

$$\begin{aligned}
\mathfrak{N}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K})) \cdot \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, 2\varepsilon)\Big] \geq& \mathfrak{N}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K})) \cdot \max_{\mathbf{h} \in \mathfrak{G}} \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{h}, \varepsilon)\Big],\\
\geq& \mathcal{L}^{\mathbf{W}}\Big[T\Big(\mathbb{B}_\alpha(\mathbf{1}, \varepsilon), \delta_\eta(\mathbf{K})\Big)\Big] \geq \Phi\Big(\eta + \Phi^{-1}\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon)\Big]\Big)\Big),
\end{aligned}$$

and taking Logarithms yields (17.1.3). □

17.1.2 Proof of Theorem 17.1.2

For this proof, we apply Proposition 17.1.3 with specific choices of ε and η.

Proof. Using Equation 17.1.2 for $\eta, \varepsilon > 0$ we have

$$\mathfrak{H}_{d_\alpha}(2\varepsilon, \delta_\eta(\mathbf{K})) \leq \tfrac{\eta^2}{2} + \mathfrak{B}(\varepsilon).$$

By the properties of the Dilation operator it follows that $\mathfrak{H}_{d_\alpha}(\varepsilon, \delta_\eta(\mathbf{K})) = \mathfrak{H}_{d_\alpha}(\varepsilon/\eta, \mathbf{K})$. Making the substitution $\eta = \sqrt{2\mathfrak{B}(\varepsilon)}$ and using that $\mathfrak{B}$ is regularly varying at infinity leads to

$$\mathfrak{H}_{d_\alpha}\Big(\tfrac{2\varepsilon}{\sqrt{2\mathfrak{B}(\varepsilon)}}, \mathbf{K}\Big) \leq 2\mathfrak{B}(\varepsilon).$$

Finally, relabeling $\varepsilon' = \frac{2\varepsilon}{\sqrt{2\mathfrak{B}(\varepsilon)}}$ which means $\varepsilon' \gtrsim \sqrt{2}\varepsilon^{\frac{\beta+1/2}{\beta}}$ and $\beta = \frac{1}{2\varrho} - \alpha$ so

$$\mathfrak{H}_{d_\alpha}\Big(\varepsilon', \mathbf{K}\Big) \leq \Big(\tfrac{2\varepsilon}{\varepsilon'}\Big)^2 \lesssim (\varepsilon')^{\frac{-1}{\frac{1}{2}+\beta}}.$$

Now additionally suppose that $\mathcal{L}^W$ satisfies Assumption 16.0.2 so that $\mathfrak{B}(\varepsilon) \gtrsim \varepsilon^{\frac{-1}{\beta}}$. For the second inequality, for $\eta, \varepsilon > 0$, we use Equation (17.1.3) with the substitution $-\eta = \Phi^{-1}\big(\mathcal{L}^{\mathbf{W}}[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon)]\big)$.

This yields

$$\mathfrak{H}_{d_\alpha}\Big(\tfrac{\varepsilon}{\Phi^{-1}\big(\mathcal{L}^{\mathbf{W}}[\mathbb{B}_\alpha(\mathbf{1},\varepsilon)]\big)}, \mathbf{K}\Big) \geq \mathfrak{B}(2\varepsilon) + \log(1/2).$$

Next, using the known limit

$$\lim_{x\to+\infty} \frac{-\Phi^{-1}(\exp(-x^2/2))}{x} = 1,$$

we equivalently have that

$$\frac{\Phi^{-1}\Big(\exp\big(-\mathfrak{B}(\varepsilon)\big)\Big)^2}{2} \sim \mathfrak{B}(\varepsilon),$$

as $\varepsilon \to 0$ since $\mathfrak{B}(\varepsilon) \to 0$. From here we conclude that as $\varepsilon \searrow 0$ we have

$$\frac{\varepsilon}{\Phi^{-1}\big(\mathcal{L}^{\mathbf{W}}[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon)]\big)} \sim \frac{\varepsilon}{\sqrt{2\mathfrak{B}(\varepsilon)}}.$$

Therefore, for ε small enough and using that $\mathfrak{B}$ varies regularly, we obtain

$$\mathfrak{H}_{d_\alpha}\Big(\tfrac{\varepsilon}{\sqrt{2\mathfrak{B}(\varepsilon)}}, \mathbf{K}\Big) \gtrsim \mathfrak{B}(2\varepsilon) + \log(1/2) \gtrsim \frac{\mathfrak{B}(\varepsilon)}{2}.$$

We conclude by making the substitution $\varepsilon' = \frac{\varepsilon}{\sqrt{2\mathfrak{B}(\varepsilon)}}$. □

17.2 Optimal Quantization and Empirical Distributions

In this section, we prove the link between Metric Entropy and Optimal Quantization and solve the asymptotic rate of convergence for the quantization problem of a Gaussian rough path.

17.2.1 Optimal Quantization

This section follows the ideals of [GLP03], although a similar result proved using a different method can be found in [DFMS03].

Theorem 17.2.1. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 and let $\mathcal{L}^{\mathbf{W}}$ be the law of the lift to the Gaussian rough path. Then for any $1 \leq r < \infty$*

$$\mathfrak{B}^{-1}\Big(\log(n)\Big) \lesssim \mathfrak{E}_{n,r}(\mathcal{L}^{\mathbf{W}}) \tag{17.2.1}$$

where $\mathfrak{B}$ is the Small Ball Probability of the measure $\mathcal{L}^{\mathbf{W}}$.

Proof. Using Definition B.0.3, we have

$$\begin{aligned}
\mathfrak{E}_{n,r}(\mathcal{L}^{\mathbf{W}})^r \geq& \int_{\Big(\cup_{\mathfrak{c}\in\mathfrak{C}_n}\mathbb{B}_\alpha\Big(\mathfrak{c},\mathfrak{B}^{-1}(\log(2n))\Big)\Big)^c} \min_{\mathfrak{c}\in\mathfrak{C}_n} d_\alpha(\mathbf{X},\mathfrak{c})^r d\mathcal{L}^{\mathbf{W}}(\mathbf{X}),\\
\geq& \mathfrak{B}^{-1}\Big(\log(2n)\Big)^r \left(1 - \mathcal{L}^{\mathbf{W}}\Big[\bigcup_{\mathfrak{c}\in\mathfrak{C}_n} \mathbb{B}_\alpha\Big(\mathfrak{c},\mathfrak{B}^{-1}(\log(2n))\Big)\Big]\right),\\
\geq& \mathfrak{B}^{-1}\Big(\log(2n)\Big)^r \Big(1 - n\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha\Big(\mathbf{1},\mathfrak{B}^{-1}(\log(2n))\Big)\Big]\Big) \geq \frac{\mathfrak{B}^{-1}\Big(\log(2n)\Big)^r}{2}
\end{aligned}$$

by applying Lemma 15.1.1. □

17.2.2 Convergence of Empirical Measure

We now turn our attention to the problem of sampling and the rate of convergence of Empirical measures. In general, the quantization problem is only theoretical as obtaining the codebook and partition that attain the minimal quantization error is computationally more complex than beneficial. An Empirical distribution removes this challenge at the sacrifice of optimality and the low probability event that the approximation will be far in the Wasserstein distance from the true distribution.

Definition 17.2.2. *For enhanced Gaussian measure $\mathcal{L}^{\mathbf{W}}$, let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space containing n independent, identically distributed enhanced Gaussian random variables $(\mathbf{W}^i)_{i=1,\dots,n}$. Let $\mathfrak{s}_i$ be the Voronoi partition of $G\Omega_\alpha(\mathbb{R}^d)$*

$$\mathfrak{s}_i := \Big\{\mathbf{X} \in G\Omega_\alpha(\mathbb{R}^d); d_\alpha(\mathbf{X},\mathbf{W}^i) = \min_{j=1,\dots,n} d_\alpha(\mathbf{X},\mathbf{W}^i)\Big\}.$$

Then we define the weighted empirical measure to be the random variable $\mathcal{M}:\Omega\to\mathcal{P}_2\big(G\Omega_\alpha(\mathbb{R}^d)\big)$

$$\mathcal{M}_n = \sum_{i=1}^n \mathcal{L}^{\mathbf{W}}(\mathfrak{s}_i)\delta_{\mathbf{W}^i}. \tag{17.2.2}$$

Note that the quantities $\mathcal{L}^{\mathbf{W}}(\mathfrak{s}_i)$ are random and $\sum_{i=1}^n \mathcal{L}^{\mathbf{W}}(\mathfrak{s}_i) = 1$. The weights are in general NOT uniform. We think of $\mathcal{M}_n$ as a (random) approximation of the measure $\mathcal{L}^{\mathbf{W}}$ and in this section we study the random variable $\mathbb{W}^{(2)}(\mathcal{M}_n,\mathcal{L}^{\mathbf{W}})$ and its mean square convergence to 0 as $n\to\infty$.

This next Theorem is an adaption of the method found in [DFMS03].

Theorem 17.2.3. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 and let $\mathcal{L}^{\mathbf{W}}$ be the law of the lift to the Gaussian rough path. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space containing a sequence of independent, identically distributed Gaussian rough paths with law $\mathcal{L}^{\mathbf{W}}$. Let $\mathcal{M}_n$ be the Empirical measure with samples drawn from the measure $\mathcal{L}^{\mathbf{W}}$. Then for any $1 \leq r < \infty$*

$$\mathbb{E}\Big[\mathbb{W}^{(r)}_{d_\alpha}\Big(\mathcal{L}^{\mathbf{W}},\mathcal{M}_n\Big)^r\Big]^{1/r} \lesssim \mathfrak{B}^{-1}\Big(\log(n)\Big). \tag{17.2.3}$$

where $\mathfrak{B}$ is the Small Ball Probability of the measure $\mathcal{L}^{\mathbf{W}}$.

Proof. By the definition of the Wasserstein distance, we have

$$\begin{aligned}\mathbb{W}^{(r)}_{d_\alpha}\Big(\mathcal{L}^{\mathbf{W}}, \mathcal{M}_n\Big)^r &= \inf_{\gamma\in\mathcal{P}\big(G\Omega_\alpha(\mathbb{R}^d)\times G\Omega_\alpha(\mathbb{R}^d)\big)} \int_{G\Omega_\alpha(\mathbb{R}^d)\times G\Omega_\alpha(\mathbb{R}^d)} d_\alpha(\mathbf{X},\mathbf{Y})^r \gamma(d\mathbf{X}, d\mathbf{Y})\\ &\le \int_{G\Omega_\alpha(\mathbb{R}^d)} \min_{j=1,\dots,n} d_\alpha\Big(\mathbf{X}, \mathbf{W}^j\Big)^r d\mathcal{L}^{\mathbf{W}}(\mathbf{X})\end{aligned}$$

Thus taking expectations, we have

$$\begin{aligned}&\mathbb{E}\left[\mathbb{W}^{(r)}_{d_\alpha}\Big(\mathcal{L}^{\mathbf{W}}, \mathcal{M}_n\Big)^r\right]\\ &= \int_{G\Omega_\alpha(\mathbb{R}^d)^{\times n}} \left(\int_{G\Omega_\alpha(\mathbb{R}^d)} \min_{j=1,\dots,n} d_\alpha\Big(\mathbf{X}, \mathbf{W}^j\Big)^r d\mathcal{L}^{\mathbf{W}}(\mathbf{X})\right) d\left(\mathcal{L}^{\mathbf{W}}\right)^{\times n}(\mathbf{W}^1, ..., \mathbf{W}^n)\end{aligned}$$

A change in the order of integration yields

$$\mathbb{E}\left[\mathbb{W}^{(r)}_{d_\alpha}\Big(\mathcal{L}^{\mathbf{W}}, \mathcal{M}_n\Big)^r\right] = 2^r \int_0^\infty \int_{G\Omega_\alpha(\mathbb{R}^d)} \Big(1 - \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{X}, 2\varepsilon^{1/r})\Big]\Big)^n d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) d\varepsilon. \tag{17.2.4}$$

Firstly, choose n large enough so that for any $0 < \varepsilon < 1$, $\sqrt{\log(n)} > \Phi^{-1}\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(1, \varepsilon^{1/r})\Big]\Big)$. Secondly, choose $c > 0$ such that $\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(1, c^{1/r})\Big] \le \Phi\Big(\frac{-1}{\sqrt{2\pi}}\Big)$.

For n and ε fixed, we label the set

$$A_{\varepsilon,n} := \left\{ \mathbf{X}\in G\Omega_\alpha(\mathbb{R}^d) : I(\mathbf{X},\varepsilon) \le \frac{\left(\frac{\sqrt{\log(n)}}{3} - \Phi^{-1}\Big(\mathcal{L}^{\mathbf{W}}\big[\mathbb{B}_\alpha(1, \varepsilon^{1/r})\big]\Big)\right)^2}{2} \right\}.$$

This can equivalently be written as

$$A_{\varepsilon,n} := \left\{ T^h[\mathbf{X}] \in G\Omega_\alpha(\mathbb{R}^d) : h \in \left(\frac{\sqrt{\log(n)}}{3} - \Phi^{-1}\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1}, \varepsilon^{1/r})\Big]\Big)\right)\mathcal{K}, \quad \mathbf{X}\in\mathbb{B}_\alpha(\mathbf{1},\varepsilon)\right\}$$

Then we divide the integral in Equation 17.2.4 into

$$\begin{aligned}\mathbb{E}\left[\mathbb{W}^{(r)}_{d_\alpha}\Big(\mathcal{L}^{\mathbf{W}}, \mathcal{M}_n\Big)^r\right] =& 2^r \int_0^c \int_{A_{\varepsilon,n}} \Big(1 - \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{X}, 2\varepsilon^{1/r})\Big]\Big)^n d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) d\varepsilon & (17.2.5)\\ &+ 2^r \int_0^c \int_{A^c_{\varepsilon,n}} \Big(1 - \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{X}, 2\varepsilon^{1/r})\Big]\Big)^n d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) d\varepsilon & (17.2.6)\\ &+ 2^r \int_c^\infty \int_{G\Omega_\alpha(\mathbb{R}^d)} \Big(1 - \mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{X}, 2\varepsilon^{1/r})\Big]\Big)^n d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) d\varepsilon. & (17.2.7)\end{aligned}$$

Firstly, using Corollary 15.1.4

$$\begin{aligned}(17.2.5) \le& 2^r \int_0^c \int_{A_{\varepsilon,n}} \Big(1 - \exp\Big(-I(\mathbf{X},\varepsilon) - \mathfrak{B}(\varepsilon^{1/r})\Big)\Big) d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) d\varepsilon\\ \le& 2^r \int_0^c \left(1 - \exp\left(-\Big(\frac{\sqrt{\log(n)}}{3} + \sqrt{2\mathfrak{B}(\varepsilon^{1/r})}\Big)^2\right)\right)^n d\varepsilon\\ \le& 2^r \int_0^{\mathfrak{B}^{-1}\left(\frac{\log(n)}{8}\right)^r} d\varepsilon + 2^r \int_{\mathfrak{B}^{-1}\left(\frac{\log(n)}{8}\right)^r}^{\mathfrak{B}^{-1}\left(\frac{\log(n)}{8}\right)^r \vee c} \left(1 - \exp\left(-\Big(\frac{\sqrt{\log(n)}}{3} + \sqrt{2\mathfrak{B}(\varepsilon^{1/r})}\Big)^2\right)\right)^n d\varepsilon\\ \le& 2^r \mathfrak{B}^{-1}\left(\frac{\log(n)}{8}\right)^r + 2^r \left(1 - \exp\left(-\frac{25\log(n)}{36}\right)\right)^n\end{aligned}$$

since $c<1$.

Now, since $\log\left(\frac{1}{\varepsilon}\right)=o\Big(\mathfrak{B}(\varepsilon)\Big)$ as $\varepsilon\to 0$, we have that $\forall p,q$ that $\exp\left(\frac{-n}{p}\right)=o\Big(\mathfrak{B}^{-1}(n)^q\Big)$. Therefore

$$\begin{aligned}\left(1-\exp\left(-\frac{25\log(n)}{36}\right)\right)^n =&\Big(1-\frac{1}{n^{25/36}}\Big)^n\\ \leq&\exp\Big(-n^{11/36}\Big)=o\left(\exp\Big(\frac{-\log(n)}{2}\Big)\right).\end{aligned}$$

Next, applying Lemma 15.1.5 to

$$\begin{aligned}\mathcal{L}^{\mathbf{W}}\Big[A^c_{\varepsilon,n}\Big]=&1-\mathcal{L}^{\mathbf{W}}\Big[A_{\varepsilon,N}\Big]\\ \leq&1-\Phi\left(\Phi^{-1}\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1},\varepsilon^{1/r})\Big]\Big)+\frac{\sqrt{\log(n)}}{3}-\Phi^{-1}\Big(\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{1},\varepsilon^{1/r})\Big]\Big)\right)\\ =&1-\Phi\left(\frac{\sqrt{\log(n)}}{3}\right)\leq\exp\left(-\frac{\log(n)}{18}\right).\end{aligned}$$

Third and finally, we make the substitution

$$\begin{aligned}(17.2.7)\leq&2^r\int_{G\Omega_\alpha(\mathbb{R}^d)}\int_0^\infty\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}^c(\mathbf{X},2\varepsilon^{1/r})\Big]d\varepsilon\cdot\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}^c(\mathbf{X},2c^{1/r})\Big]^{n-1}d\mathcal{L}^{\mathbf{W}}(\mathbf{X})\\ \leq&\mathbb{E}\Big[\|\mathbf{W}\|_\alpha^r\Big]\int_{G\Omega_\alpha(\mathbb{R}^d)}\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}^c(\mathbf{X},2c^{1/r})\Big]^{n-1}d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) &(17.2.8)\\ &+\int_{G\Omega_\alpha(\mathbb{R}^d)}\|\mathbf{X}\|_\alpha^r\cdot\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}^c(\mathbf{X},2c^{1/r})\Big]^{n-1}d\mathcal{L}^{\mathbf{W}}(\mathbf{X}) &(17.2.9)\end{aligned}$$

to account for the integral over (c,∞). Next, we partition this integral over $A_{c,n}$ and $A^c_{c,n}$. Arguing as before, we have

$$\begin{aligned}&\int_{A_{c,n}}\Big(1-\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{X},2c^{1/r})\Big]\Big)^{n-1}d\mathcal{L}^{\mathbf{W}}(\mathbf{X})\\ &\qquad\leq\left(1-\exp\Big(-\frac{25\log(n)}{36}\Big)\right)^{n-1}\leq o\Big(\mathfrak{B}^{-1}\Big(\log(n)\Big)^q\Big)\end{aligned}$$

and

$$\begin{aligned}&\int_{A^c_{c,n}}\Big(1-\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}_\alpha(\mathbf{X},2c^{1/r})\Big]\Big)^{n-1}d\mathcal{L}^{\mathbf{W}}(\mathbf{X})\\ &\qquad\leq\sup_{0<c}\mathcal{L}^{\mathbf{W}}\Big[A^c_{c,n}\Big]\leq o\Big(\mathfrak{B}^{-1}\Big(\log(n)\Big)^q\Big).\end{aligned}$$

for any choice of q. Therefore

$$\begin{aligned}(17.2.8)\leq&o\Big(\mathfrak{B}^{-1}\Big(\log(n)\Big)^q\Big)\\ (17.2.9)\leq&\mathbb{E}\Big[\|\mathbf{W}\|_\alpha^{2r}\Big]^{1/2}\cdot\left(\int_{G\Omega_\alpha(\mathbb{R}^d)}\mathcal{L}^{\mathbf{W}}\Big[\mathbb{B}^c_\alpha(\mathbf{X},2c^{1/r})\Big]^{2(n-1)}d\mathcal{L}^{\mathbf{W}}(\mathbf{X})\right)^{1/2}\\ \leq&o\Big(\mathfrak{B}^{-1}\Big(\log(n)\Big)^q\Big)\end{aligned}$$

Thus

$$(17.2.4)\leq 2^r\mathfrak{B}^{-1}\Big(\tfrac{\log(n)}{8}\Big)^r+o\Big(\mathfrak{B}^{-1}\Big(\log(n)\Big)^r\Big).$$

□

Part VI

Concluding Remark

Chapter 18

Further Research

18.1 Malliavin Differentiability Integration by parts for McKean-Vlasov Equations

Malliavin differentiability of McKean-Vlasov equations is a simple extension of Theorem 9.3.2. See for example [CM18] and the upcoming work

Conjecture 18.1.1. *Let* $b : [0,T]\times\Omega\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ *and* $\sigma : [0,T]\times\Omega\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d'}$ *satisfy Assumption 9.3.1 with additional regularity conditions on the measure. Then the stochastic differential equation*

$$X_t = \theta + \int_0^t b(s,\omega,X_s(\omega),\mathcal{L}_s^X)ds + \int_0^t \sigma(s,\omega,X_s(\omega),\mathcal{L}_s^X)dW_s$$

is Malliavin differentiable with Malliavin derivative

$$\begin{aligned} D_sX_t(\omega) =&\sigma(s,\omega,X_s(\omega),\mathcal{L}_s^X) + \int_s^t U(s,r,\omega)dr + \int_s^t V(s,r,\omega)dW_r \\ &+ \int_s^t \nabla_x b\big(r,\omega,X_r(\omega),\mathcal{L}_s^X\big)D_sX_r(\omega)dr \\ &+ \int_s^t \nabla_x \sigma\big(r,\omega,X_r(\omega),\mathcal{L}_s^X\big)D_sX_r(\omega)dW_r. \end{aligned} \tag{18.1.1}$$

However, parametric differentiability is more involved due to the dependency of the initial condition in the law.

18.2 Support Theorem for Gaussian rough paths and probabilistic rough paths

There are a number of questions relating to support theorems that are not addressed in this thesis or in [CRS19]. These include supports for McKean-Vlasov Equations driven by a Gaussian noise other than Brownian motion and the introduction of measure dependencies in the diffusion terms.

The first problem to overcome is to construct quantizations for these Gaussian measures. In practice, this is not a lot more involved although the projective subspace needs to be constructed for each noise. In the case of Fractional Brownian motion, we obtained fractional wavelets. Next, we need to construct a partition for the quantization. One of the details that needs establishing is whether the Voronoi boundaries have 0 measure or not. Many of the methods used in Chapter 12 will work for a general noise.

Next, we need to verify that the alegbraic properties of systems of interacting equations driven by probabilistic rough paths are equivalent to those demonstrated in Chapter 13. My aspiration is to address these problems in the future.

This leads us to the expected result:

Conjecture 18.2.1. *Let $\xi \in \mathcal{P}_r(\mathbb{R}^d)$. Let $\mathbf{W}$ be the probabilistic rough path of a Gaussian process satisfying Assumption C.2.12 with law $\mathcal{L}^{\mathbf{W}}$. Let $\mathbf{q}_n$ be the sequence of quantizations that approximate $\xi \times \mathcal{L}^{\mathbf{W}}$.*

Let $\mathcal{L}^{\Phi([\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1})}$ be the law of the Interacting Particle System driven by the quantization. Suppose that $b : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d'}$ be adequately differentiable in spacial and measure variables. Then the $\mathcal{L}^{\mathbf{X}}$, the law of the McKean-Vlasov Equation, satisfies

$$\operatorname{supp}(\mathcal{L}^{\mathbf{X}}) = \bigcap_{m=1}^{\infty} \overline{\bigcup_{n\geq m} \Big\{ \Theta_{b,\sigma}(\mathcal{L}^{\Phi([\xi\times\mathcal{L}^{\mathbf{W}}]\circ\mathbf{q}_n^{-1})}, x, \mathbf{h}) : h \in \mathcal{H}, x \in \operatorname{supp}(\xi) \Big\}}^{\rho_{\alpha-H\ddot{o}l;[0,T]}}. \quad (18.2.1)$$

18.3 Convergence of Empirical Measure (non-weighted)

We could study the Empirical Measure with uniform weights. In general, this is the form that the Empirical distribution takes, although we belive that the proof rate of convergence for the Weighted Empirical distribution is easier.

Definition 18.3.1. *For enhanced Gaussian measure $\mathcal{L}^{\mathbf{W}}$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space containing n independent, identically ditributed enhanced Gaussian random variables $(\mathbf{W}^i)_{i=1,\dots,n}$.*

Then we define the Empirical Measure to be the random variable $\mathcal{E}_n : \Omega \to \mathcal{P}_2\big(G\Omega_\alpha(\mathbb{R}^d)\big)$

$$\mathcal{E}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\mathbf{W}^i}. \quad (18.3.1)$$

Conjecture 18.3.2. *Let $\mathcal{L}^W$ be a Gaussian measure satisfying Assumption C.2.12 and let $\mathcal{L}^{\mathbf{W}}$ be the law of the lift to the Gaussian rough path. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space containing a sequence of independent, identically distributed Gaussian rough paths with law $\mathcal{L}^{\mathbf{W}}$. Let $\mathcal{E}_n$ be the Empirical measure with samples drawn from the measure $\mathcal{L}^{\mathbf{W}}$. Then for any $1 \leq r < \infty$*

$$\mathbb{E}\Big[\mathbb{W}^{(r)}_{d_\alpha}\Big(\mathcal{L}^{\mathbf{W}}, \mathcal{E}_n\Big)^r\Big] \approx \mathbb{E}\Big[\mathbb{W}^{(r)}_{d_\alpha}\Big(\mathcal{L}^{\mathbf{W}}, \mathcal{M}_n\Big)^r\Big] \quad (18.3.2)$$

Proof. Following the methods of [BLG14] and [Boi11], this should be unremarkable. The contribution would be proving this result without having to resort to Talagrand inequalities. □

Part VII

Appendix

Appendix A

Miscellaneous mathematics

A.1 Pathspace analysis

Let $C^0([0,T];\mathbb{R}^d)$ be the space of continuous functions over the interval $[0,T]$ taking values in the vector space $\mathbb{R}^d$ that start at 0 paired with the supremum norm. For $\alpha \in (0,1)$, we define the α-Hölder norm

$$\|\psi\|_\alpha = \sup_{s,t\in[0,T]} \frac{|\psi(t)-\psi(s)|}{|t-s|^\alpha}.$$

Let $C^\alpha([0,T];\mathbb{R}^d)$ be the subset of $C^0([0,T];\mathbb{R}^d)$ such that $\|\cdot\|_\alpha$ is finite. For $\alpha<\beta<1$, β-Hölder continuous paths are compactly embedded in the space of α-Hölder continuous paths e.g. the spaces $C^\beta([0,T];\mathbb{R}^{d'}) \Subset C^\alpha([0,T];\mathbb{R}^{d'})$. Although the space $C^\alpha([0,T];\mathbb{R}^{d'})$ is not separable, the subset $C^{\alpha,0}([0,T];\mathbb{R}^{d'}) := \overline{C^\beta([0,T];\mathbb{R}^{d'})}^{\alpha-\text{Hölder}}$ is separable.

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space carrying a d'-dimensional Brownian Motion on the interval $[0,T]$ where throughout $T>0$. The Filtration on this space satisfies the usual assumptions. We denote by $\mathbb{E}$ and $\mathbb{E}[\cdot|\mathcal{F}_t]$ the usual expectation and conditional expectation operator (with respect to $\mathbb{P}$) respectively. For a random variable X we denote its probability distribution (or Law) by $\mathcal{L}^X$; the law of a process $(Y_t)_{t\in[0,T]}$ at time t is denoted by $\mathcal{L}_t^Y$.

For μ, a probability measure on $(E,\mathcal{B})$, we define the support of μ, denoted $\operatorname{supp}(\mu)$, to be the set of points $x\in E$ such that every open neighbourhood of x has positive measure. Equivalently, it is the smallest closed set of full measure.

Let E be a separable Banach space. Then it is well known that the Borel σ-algebra and the cylindrical σ-algebra are the same (see for example [Bog98]). Let $\mathcal{H}$ be the Reproducing Kernel Hilbert Space (RKHS) of the Gaussian measure. We denote the unit ball in the RKHS norm as $\mathcal{K}$. It is well known that the set $\mathcal{K}$ is compact in the Banach space topology of E and $\mathcal{H}$ is dense in the support of $\mathcal{L}$.

We consider the law of a Gaussian process as a measure on pathspace, that is a measure over the space of continuous paths starting at $0\in\mathbb{R}^{d'}$. We are interested in the space of α-Hölder continuous paths for $\alpha<\frac{1}{2\varrho}$ and the topology induced by this norm where $\varrho\in[1,3/2)$. For any choice of $\alpha<\frac{1}{2\varrho}$, we can find $\alpha<\alpha'<\frac{1}{2\varrho}$ for which the Gaussian process will be α'-Hölder continuous. Therefore, we will always have that the Gaussian process takes values in $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$ and we do not concern ourselves with separability further.

Definition A.1.1 (Haar Functions). *Let $t\in[0,T]$. For $p\in\mathbb{N}_0$ and $m\in\{1,...,2^p\}$, define the sequence of values $t_{pm}^0=\frac{(m-1)T}{2^p}$, $t_{pm}^1=\frac{(2m-1)T}{2^{p+1}}$ and $t_{pm}^2=\frac{mT}{2^p}$. Define the functions $H_{00}(t)=1$ and*

$$H_{pm}(t)=\begin{cases}\sqrt{\frac{2^p}{T}}, & \text{if } t\in[t_{pm}^0,t_{pm}^1),\\ -\sqrt{\frac{2^p}{T}}, & \text{if } t\in[t_{pm}^1,t_{pm}^2),\\ 0, & \text{otherwise}.\end{cases}$$

These are called the Haar functions, a orthonormal collection of functions in $L^2([0,T];\mathbb{R})$.

The Schauder function are similarly defined $G_{pm}(t)=\int_0^t H_{pm}(s)ds$.

The Haar functions form an orthonormal basis on the space $L^2([0,T];\mathbb{R})$ with the canonical inner product. Therefore, we define the Fourier coefficients $\psi_{pm} = \int_0^T H_{pm}(s)\psi(s)ds$ and the set

$$\Lambda := \Big\{(p,m) : p \in \mathbb{N}_0, m \in \{1,...,2^p\}\Big\} \cup \Big\{(0,0)\Big\}.$$

We do not include the pair $(p,m) = (-1,0)$ as throughout we will be dealing with Gaussian processes which are 0 at $t = 0$.

Next, for some continuous path ψ taking values in $\mathbb{R}^{d'}$, we define the Schauder Fourier coefficients to be

$$\psi_{pm} := \langle H_{pm}, \mathrm{d}\psi\rangle := \sqrt{\tfrac{2^p}{T}}\left[2\psi(t^1_{pm}) - \psi(t^0_{pm}) - \psi(t^2_{pm})\right] \in \mathbb{R}^{d'}, \quad \text{for} \quad (p,m) \in \Lambda; \tag{A.1.1}$$

additionally $\psi_{00} := \langle H_{00}, \mathrm{d}\psi\rangle = \psi(1) - \psi(0)$. Let us denote $\Lambda_N = \{(p,m) \in \Lambda : p \leq N\}$ as a truncation of Λ.

The following Theorem, often referred to as the Cielsielski Isomorphism, provides the link between wavelet theory and rough paths.

Theorem A.1.2 ([HIPP14])**.** *For $\alpha > 0$, let $\|\cdot\|_\alpha$ be the α-Hölder norm. Let $\psi \in C^0([0,T];\mathbb{R}^{d'})$. We have that $\|\cdot\|_\alpha$ is equivalent to*

$$\|\psi\|'_\alpha = \sup_{(p,m)\in\Lambda} 2^{(\alpha-1/2)p}|\psi_{pm}|. \tag{A.1.2}$$

If, in addition, we have that

$$\lim_{p\to\infty} 2^{p(\alpha-1/2)} \sup_{1\leq m\leq 2^p} |\psi_{pm}| = 0$$

we say that $\psi \in C^{\alpha,0}([0,T];\mathbb{R}^{d'})$. This space is a separable subset of $C^\alpha([0,T];\mathbb{R}^{d'})$.

Example A.1.3 (Cielsielski Representation of Brownian motion)**.** *Due to the orthogonality of the Schauder functions in the RKHS of Brownian motion, we can represent Brownian motion as*

$$W_t = \sum_{(p,m)\in\Lambda} W_{pm}G_{pm}(t) \quad t \in [0,T] \tag{A.1.3}$$

where W_{pm} is a sequence of d'-dimensional, independent, standard normally distributed random variables. Thus

$$\|W\|_\alpha = \sup_{(p,m)\in\Lambda} 2^{p(\alpha-1/2)}|W_{pm}|.$$

A.2 Malliavin Calculus

We briefly summarise some standard results relating to Gaussian processes and Gaussian measures.

Definition A.2.1. *A centred Gaussian measure $\mathcal{L}$ on a real separable Banach space E equipped with its Borel σ-algebra $\mathcal{B}$ is a Borel probability measure on $(E,\mathcal{B})$ such that the law of each continuous linear functional on E is Gaussian with mean 0.*

A.2.1 Chaos expansions and Malliavin Calculus

Lemma A.2.2. *Let θ be $\mathcal{F}_0$-measurable. Then $D_s\theta = 0$ for any $s \in (0,T]$.*

Proof. We have that θ is $\mathcal{F}_0$-measurable, so it must also be $\mathcal{F}_t$-measurable for any $t \geq 0$. This means it will have an Itô-Wiener Chaos expansion and we denote this by

$$\theta = \theta_0 + \sum_{n=1}^{\infty} I_n[\theta_n] \quad \text{where} \quad \theta_n \in L^2([0,T]^n).$$

Next, using that for any choice of $f \in L^2([0,T]^n)$

$$\mathbb{E}\Big[\theta \cdot I_n[f]\Big|\mathcal{F}_0\Big] = \theta \cdot \mathbb{E}\Big[I_n[f]\Big|\mathcal{F}_0\Big] = 0,$$

we conclude that $\theta_n = 0$ for $n \geq 1$. Finally, using the formula from [Nua06] for the chaos expansion of the Malliavin derivative, we conclude that $D_s\theta = 0$ □

A.2.2 Classical results on the Cameron Martin transforms

We recall two useful results from [ÜZ00]. First we introduce the notation for a Doléans-Dade exponential over $[0,T]$ of some sufficient integrable $\mathbb{R}^{d'}$-valued process, $(M_t)_{t\in[0,T]}$, namely, we define for $t \in [0,T]$ and an d'-dimensional Brownian motion W,

$$\mathcal{E}(M)_t = \exp\Big(\int_0^t M_s dW(s) - \frac{1}{2}\int_0^t |M_s|^2 ds\Big). \tag{A.2.1}$$

Proposition A.2.3 (The Cameron-Martin Formula – [ÜZ00])**.** *Let F be an $\mathcal{F}_T$-measurable random variable. For $h \in \mathcal{H}$ let $\mathcal{E}(\dot{h})$ be the associated Doléans-Dade exponential.*

Then, when both sides are well defined,

$$\mathbb{E}\big[\,F(\omega+h)\,\big] = \mathbb{E}\Big[F\exp\Big(\int_0^T \dot{h}(s)dW(s) - \frac{1}{2}\int_0^T |\dot{h}(s)|^2 ds\Big)\Big] = \mathbb{E}\big[\,F(\omega)\mathcal{E}(\dot{h})_T\,\big].$$

Moreover, $\forall h \in \mathcal{H}$ and $\forall p \geq 1$ that $\mathcal{E}(\dot{h}) \in \mathcal{S}^p([0,T])$.

Proposition A.2.4 (Continuity of the Cameron Martin Transform – [ÜZ00])**.** *The map τ_h : $[0,1] \to L^0(\Omega)$ defined by $t \mapsto f(\omega+th)$ is continuous map from a compact interval of the real line to a measurable function with respect to the topology of convergence in probability.*

A.2.3 Malliavin Calculus for SDEs

The following result is an adaption of the proof from [Nua06, Theorem 2.2.1]

Theorem A.2.5. *Let $\theta : \Omega \to \mathbb{R}^d$, let $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \times\mathbb{R}^d \to \mathbb{R}^{d\times d'}$ be progressively measurable maps, and $L > 0$ such that*

1. *$\theta \in L^p(\mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$,*
2. *b and σ are integrable in the sense that*
$$\int_0^T |b(t,\omega,0)|dt, \quad \int_0^T |\sigma(t,\omega,0)|^2 dt < \infty,$$
3. *$\exists L > 0$ such that for almost all $s \in [0,T]$ and $\forall x, y \in \mathbb{R}^d$ we have*
$$|b(s,x) - b(s,y)|, |\sigma(s,x) - \sigma(s,y)| < L|x-y|,$$

Then there exists a unique solution to the SDE

$$X_t = \theta + \int_0^t b(r,X_r)dr + \int_0^t \sigma(r,X_r)dW_r \tag{A.2.2}$$

in the space $\mathcal{S}^\infty([0,T])$ and the solution is Malliavin differentiable. Further, there exist processes $B : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ and $\Sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d'\times d}$ such that the Malliavin Derivative satisfies the linear SDE

$$D_s X_t = \mathbb{1}_{(0,t)}(s)\Big(\sigma(s,X_s) + \int_s^t B(r,X_r)D_sX_r dr + \int_s^t \Sigma(r,X_r)D_sX_r dW_r\Big). \tag{A.2.3}$$

Appendix B

Quantization

Quantization is the problem of finding a finite support measure that approximates a given measure. We provide a brief introduction to the field of quantization. For further details, see [GL00].

Definition B.0.1. *Let $\mathcal{L}$ be a measure on a separable Banach space E endowed with the Borel σ-algebra such that $\mathcal{L} \in \mathcal{P}_2(E)$ and all sets of codimension 1 have null $\mathcal{L}$-measure.*

Let I be a countable index, let $\mathfrak{S} := \{\mathfrak{s}_i, i \in I\}$ be a partition of E and let $\mathfrak{C} := \{\mathfrak{c}_i \in E, i \in I\}$ be a codebook. For any partition $\mathfrak{S}$ and codebook $\mathfrak{C}$, we define a quantization $q : E \to E$ by

$$q(x) = \mathfrak{c}_i \quad \text{for } x \in \mathfrak{s}_i, \qquad q(E) = \mathfrak{C}$$

so that

$$\mathcal{L} \circ q^{-1}(\cdot) = \sum_{i \in I} \mathcal{L}(\mathfrak{s}_i)\delta_{\mathfrak{c}_i}(\cdot) \in \mathcal{P}_2(E).$$

The collection of all quantizations is denoted $\mathfrak{Q}$.

Definition B.0.2. *Let $\mathfrak{P} \subset [0,1]^{\mathbb{N}}$ be the set of probability vectors e.g. for every $\mathfrak{p} = (\mathfrak{p}_i)_{i\in\mathbb{N}}$, we have $\mathfrak{p}_i \in [0,1]$ and $\sum_{i\in\mathbb{N}} \mathfrak{p}_i = 1$.*

Given a partition $\mathfrak{S}$ of a space E, we have that the sequence $(\mathcal{L}(\mathfrak{s}_i))_{\mathfrak{s}_i\in\mathfrak{S}}$ is a probability vector.

Definition B.0.3 (Optimal Quantizers)**.** *Let $n \in \mathbb{N}$ and $r \in [1, \infty)$. The minimal n^{th} quantization error of order r of a measure $\mathcal{L}$ on a separable Banach space E is defined to be*

$$\mathfrak{E}_{n,r}(\mathcal{L}) = \inf \left\{ \Big(\int_E \min_{\mathfrak{c}\in\mathfrak{C}} \|x - \mathfrak{c}\|_E^r d\mathcal{L}(x) \Big)^{\frac{1}{r}} : \mathfrak{C} \subset E, 1 \leq |\mathfrak{C}| \leq n \right\}.$$

A Codebook $\mathfrak{C} = \{\mathfrak{c}_i, i \in I\}$ with $1 \leq |\mathfrak{C}| \leq n$ is called an n-optimal set of centres of $\mathcal{L}$ (of order r) if

$$\mathfrak{E}_{n,r}(\mathcal{L}) = \Big(\int_E \min_{i=1,\dots,n} \|x - \mathfrak{c}_i\|_E^r d\mathcal{L}(x) \Big)^{\frac{1}{r}}$$

Given a finite collection of elements $(\mathfrak{c}_i)_{i=1,\dots,n}$, the optimal way to choose the partition of E is to use the nearest neighbour rule which corresponds to the Voronoi partition

$$\mathfrak{s}\Big(\mathfrak{c}_i \Big| (\mathfrak{c}_j)_{j=1,\dots,n}\Big) := \Big\{ x \in E : \|x - \mathfrak{c}_i\| = \min_{j=1,\dots,n} \|x - \mathfrak{c}_j\| \Big\} \tag{B.0.1}$$

provided the boundary of the Voronoi sets has measure 0. Sets of the form (B.0.1) are called *Voronoi sets*. Similarly, given a finite partition $(\mathfrak{s}_i)_{i=1,\dots,n}$ of E, the optimal choice of codebook is the centres of mass for the sets $\mathfrak{s}_i$ with respect to the measure $\mathcal{L}$. For brevity of notation, we write $\mathfrak{E}_n := \mathfrak{E}_{n,2}$.

B.1 Stationary Quantization

A Stationary set is a codebook with a special property: the Voronoi sets generated by codebook have barycentres equal to the codebook.

Definition B.1.1. *Let E be a separable Banach space with borel σ-algebra $\mathcal{B}$, let $n \in \mathbb{N}$ and let $\mathcal{L}$ be a measure on $(E, \mathcal{B})$ such that and all subsets of codimension 1 have null $\mathcal{L}$-measure. Let $\mathfrak{C} \subset E$ satisfy $|\mathfrak{C}| = n$*

Suppose that the Voronoi partition $\mathfrak{S}$ of E generated by the elements of $\mathfrak{C}$, containing the collection of sets $\mathfrak{s}_i := \{y \in E : \min_{j=1,...,n} \|y - \mathfrak{c}_j\| = \|y - \mathfrak{c}_i\|\}$ satisfies that

$$\frac{1}{\mathcal{L}(\mathfrak{s}_i)} \int_{\mathfrak{s}_i} y d\mathcal{L}(y) = \mathfrak{c}_i.$$

*Then we call the codebook $\mathfrak{C}$ an n-*stationary set *of the law $\mathcal{L}$.*

Theorem B.1.2 ([Lal10, Theorem 2.1])**.** *Let E be a reflexive, separable Banach space and let $\mathcal{L}$ be a measure on $(E, \mathcal{B})$. For $\mathfrak{c}_i \in E$, define $\mathfrak{A} : E^n \to \mathbb{R}$ by*

$$\mathfrak{A}(\mathfrak{c}_1, ..., \mathfrak{c}_n) = \int_E \min_{i=1,...,n} \|y - \mathfrak{c}_i\|_E^2 d\mathcal{L}(y)$$

e.g. $\mathfrak{A}(\mathfrak{c}_1, ..., \mathfrak{c}_n)$ is the mean square error between the measure $\mathcal{L}$ and the quantization with codebook $\{\mathfrak{c}_1, ..., \mathfrak{c}_n\}$ and partition equal to the Voronoi sets of the codebook.

Then $\mathfrak{A}$ admits at least one minimum, and so an n-stationary set exists.

Remark B.1.3. *The proof of the above result relies on the Assumption that the Banach space E is reflexive. In particular, for a non-reflexive space the unit ball will be weak-$*$ compact but not weak compact (see* [FHH+01, Theorem 3.31]*). By contrast, the functional $\mathfrak{A}$ can be shown to be weak lower semicontinuous but the proof does not extend to weak-$*$ lower semicontinuity.*

In particular, we are interested in Gaussian measures over the Banach space $C^{\alpha,0}([0,T]; \mathbb{R}^{d'})$, which is not reflexive and so Theorem B.1.2 does not apply.

Lastly, it is not clear whether a stationary quantization exists in general.

Lemma B.1.4. *Let $\mathcal{L}$ be a centred Gaussian measure taking values on the Banach space E and suppose that an n-stationary set exists. Let $\mathfrak{C}$ be an n-stationary set. Then $\mathfrak{C} \subset \mathcal{H}$.*

Proof. This proof is based on a similar argument first presented in [LP02] which focuses solely on Hilbert spaces. Using that the n-stationary set exists, we have that for any $\mathfrak{c} \in \mathcal{C}$

$$\mathfrak{c} = \int_{\mathfrak{s}} x d\mathcal{L}(x) = \int_E x \cdot \frac{\mathbb{1}_{\mathfrak{s}}(x)}{\mathcal{L}(\mathfrak{s})} d\mathcal{L}(x)$$

Next, we use that $\frac{\mathbb{1}_{\mathfrak{s}}}{\mathcal{L}(\mathfrak{s})}$ is a square integrable function with respect to $\mathcal{L}$ on E and use Definition of the RKHS to conclude that the right hand side of this equation must be an element of $\mathcal{H}$. Therefore $\mathfrak{c} \in \mathcal{H}$ $\square$

Remark B.1.5. *In particular, if $q_n(W)$ denotes the quantized random variable W, then the Stationary quantization has the property that*

$$q_n(W) = \mathbb{E}[W|\mathcal{F}_n],$$

where $\mathcal{F}_n$ is the σ-algebra generated by the partition of q_n. This is a particularly useful property when it comes to establishing uniform integrability of quantizations due to the following simple argument:

Let ϕ be a convex function on a Banach space E. Then

$$\sup_{n\in\mathbb{N}} \mathbb{E}\Big[\phi(q_n(W))\Big] = \sup_{n\in\mathbb{N}} \mathbb{E}\Big[\phi\big(\mathbb{E}[W|\mathcal{F}_n]\big)\Big] \le \sup_{n\in\mathbb{N}} \mathbb{E}\Big[\mathbb{E}\big[\phi(W)|\mathcal{F}_n\big]\Big] = \mathbb{E}\Big[\phi(W)\Big]. \tag{B.1.1}$$

Lemma B.1.6. *Let $\mathcal{L}$ be a non-degenerate Gaussian measure over E with RKHS $\mathcal{H}$. Let U be a finite dimensional subspace of $\mathcal{H}$ and let P_U be the orthogonal projection operator from $\mathcal{H}$ to U extended to $E = \overline{\mathcal{H}}^E$. Then $\forall r > 1$*

$$\mathfrak{E}_{n,r}(\mathcal{L}) \lesssim \mathfrak{E}_{n,r}\big(\mathcal{L} \circ (P_U)^{-1}\big) + \left(\int_E \big\|x - P_U[x]\big\|_E^r d\mathcal{L}(x) \right)^{1/r}.$$

In particular, when the measure $\mathcal{L}$ is in some sense "concentrated" on a finite dimensional linear subspace of the Banach space E, then the quantization problem can be simplified to a finite dimensional problem.

Proof. From Definition B.0.3, we have

$$\begin{aligned}
\mathfrak{E}_{n,r}(\mathcal{L}) &= \min_{h_1,\dots,h_n \in E} \left(\int_E \min_{i=1,\dots,n} \big\|x - h_i\big\|_E^r d\mathcal{L}(x) \right)^{1/r} \\
&\le \min_{h_1,\dots,h_n \in P_U[E]} \left(\int_E \min_{i=1,\dots,n} \big\|x - h_i\big\|_E^r d\mathcal{L}(x) \right)^{1/r}, \\
&\le 2^{(r-1)/r} \min_{h_1,\dots,h_n \in P_U[E]} \left(\int\int_{P_U[E]\times(I-P_U)[E]} \min_{i=1,\dots,n} \big\|P_U[x] - h_i\big\|_E^r d\mathcal{L}\Big(P_U[x]\Big) d\mathcal{L}\Big((I-P_U)[x]\Big) \right. \\
&\qquad \left. + \int_E \big\|(I-P_U)[x]\big\|_E^r d\mathcal{L}(x) \right)^{1/r},
\end{aligned}$$

since by the assumption that P_U is a projection on $\mathcal{H}$ (rather than E), the two laws $\mathcal{L} \circ (P_U)^{-1}$ and $\mathcal{L} \circ (I - P_U)^{-1}$ are independent with respect to the joint law $\mathcal{L}$. Exploiting this, we get

$$\begin{aligned}
\mathfrak{E}_{n,r}(\mathcal{L}) \le& 2^{(r-1)/r} \left(\int_{(I-P_U)[E]} \mathfrak{E}_{n,r}(\mathcal{L} \circ (P_U)^{-1})^r d\mathcal{L}\Big((I-P_U)[x]\Big) + \int_E \big\|x\big\|_E^r d\mathcal{L}\Big((I-P_U)[x]\Big) \right)^{1/r} \\
\le& 2^{(r-1)/r} \left(\mathfrak{E}_{n,r}(\mathcal{L} \circ (P_U)^{-1}) + \Big(\int_E \big\|x - P_U[x]\big\|_E^r d\mathcal{L}(x) \Big)^{1/r} \right).
\end{aligned}$$

□

B.2 Rate of Convergence for Quantization

In the finite dimensional setting, the minimal quantization error is well understood (see [GL00]). Let $\mathcal{L}$ be a measure over a d-dimensional vector space. Then

$$\mathfrak{E}_{n,r}(\mathcal{L}) \approx n^{1/d}. \tag{B.2.1}$$

However, for a Gaussian measure over a Banach space E, the limit $d \to \infty$ is no longer meaningful. In both [DFMS03] and [GLP03], the authors investigate the relation between the minimal quantization error and the probabilities of small balls.

Theorem B.2.1 ([DFMS03], [GLP03])**.** *Let $\mathcal{L}^W$ be a Gaussian measure over a Banach space E. Let $\mathfrak{B}_W$ be the small ball probability of $\mathcal{L}^W$ defined by $\mathfrak{B}_W(\varepsilon) := -\log \mathcal{L}\big[\{x \in E : \|x\|_E < \varepsilon\}\big]$. Then for any choice of $r \ge 1$*

$$\mathfrak{E}_{n,r}(\mathcal{L}^W) \approx (\mathfrak{B}_W)^{-1}\Big(\log(n)\Big)$$

as $n \to \infty$. In particular, let $\mathcal{L}^W$ be the law of Brownian motion over $C^{\alpha,0}([0,T];\mathbb{R}^{d'})$. Then by the results of [BR92]

$$\mathfrak{E}_{n,r}(\mathcal{L}^W) \approx d'\Big(\log(n^{1/d'})\Big)^{\alpha - 1/2}. \tag{B.2.2}$$

In particular, Equation (B.2.2) provides us with a lower bound that the error of the quantization for Brownian motion cannot outstrip. However, as already explained in Remark B.1.3, there may not exist a stationary quantization that attains $\mathfrak{E}_{n,r}(\mathcal{L}^W)$.

A remarkable aspect of [DFMS03] is that the authors additionally prove that the mean square error between an empirical measure and the true Gaussian measure in the Wasserstein distance converges at the same rate as the optimal quantization error.

Appendix C

Rough Differential Equations

The theory of rough paths, first developed by Lyons in [Lyo98], is a collection of widely used and powerful tools developed to give analytic and algebraic meaning to ways in which "rough" noises drive systems of differential equations. There are several comprehensive monographs, see [FV10b, FH14, LQ02, LCL07], with each providing their own unique approach to the introduction of core material.

C.1 Algebraic Material

Let $\mathscr{H}$ be a locally finite graded connected Hopf algebra with associative product $\odot : \mathscr{H} \otimes \mathscr{H} \to \mathscr{H}$, coassociative coproduct $\Delta : \mathscr{H} \to \mathscr{H} \otimes \mathscr{H}$, unit $\mathbf{1}$ and counit $\varepsilon \in \mathscr{H}^*$, and antipode $A : \mathscr{H} \to \mathscr{H}$ such that

$$\odot \circ \Big(\mathrm{id} \otimes A\Big)[\Delta[x]] = \varepsilon(x)\mathbf{1} = \odot \circ \Big(A \otimes \mathrm{id}\Big)[\Delta[x]].$$

Thus $\mathscr{H}$ has the representation

$$\mathscr{H} = \bigoplus_{n=0}^{\infty} \mathscr{H}_{(n)}$$

where $\mathscr{H}_{(0)} = \mathbf{1}\mathbb{R}$, the vector spaces $\mathscr{H}_{(n)}$ is finite dimensional and

$$\odot : \mathscr{H}_{(n)} \otimes \mathscr{H}_{(m)} \to \mathscr{H}_{(n+m)}, \quad \Delta : \mathscr{H}_{(n)} : \bigoplus_{p+q=n} \mathscr{H}_{(p)} \otimes \mathscr{H}_{(q)}$$

Definition C.1.1. *A character on $\mathscr{H}$ is a functional $g \in \mathscr{H}^*$ such that*

$$\langle g, h_1 \odot h_2 \rangle = \langle g, h_1 \rangle \langle g, h_2 \rangle$$

for all $h_1, h_2 \in \mathscr{H}$. We call G the set of all characters on $\mathscr{H}$.

A derivation on $\mathscr{H}$ is a functional $p \in \mathscr{H}^*$ such that

$$\langle \mathfrak{l}, h_1 \odot h_2 \rangle = \langle \mathfrak{l}, h_1 \rangle \langle \varepsilon, h_2 \rangle + \langle \mathfrak{l}, h_2 \rangle \langle \varepsilon, h_1 \rangle.$$

We call $\mathfrak{L}$ the set of all derivations of $\mathscr{H}$.

It is well known that the characters of a Hopf algebra form a group with unit ε and inverse obtained by composition with the antipode. Similarly, the space of derivations forms a Lie algebra with Lie brackets defined by

$$[\mathfrak{l}_1, \mathfrak{l}_2]_{\boxtimes} = \mathfrak{l}_1 \boxtimes \mathfrak{l}_2 - \mathfrak{l}_2 \boxtimes \mathfrak{l}_1$$

and $\boxtimes$ is the product on $\mathscr{H}^*$ dual to the coproduct Δ. Further, there exist bijective diffeomorphisms between $\mathfrak{L}$ and G called the exponential map $\exp_{\boxtimes} : \mathfrak{L} \to G$ and logarithm map

$\log_\boxtimes : G \to \mathfrak{L}$ defined by

$$\exp_\boxtimes(\mathfrak{l}) := \sum_{n=0}^{\infty} \frac{\mathfrak{l}^{\boxtimes n}}{n!}, \quad \log_\boxtimes(g) := \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(g-\varepsilon)^{\boxtimes n}}{n}.$$

By inductively defining the sequence of subspaces

$$V_1 = \mathscr{H}^*_{(1)}, \quad V_{i+1} = [V_i, V_1]_\boxtimes \subset \mathscr{H}^*_{(k+1)}$$

we can represent $\mathfrak{L}$ as the completion

$$\mathfrak{L} = \prod_{i=1}^{\infty} V_i$$

of the graded Lie algebra $V_1 \oplus V_2 ...$ with respect to the natural filtration induced by the grading. Thus it also admits a filtration

$$\mathfrak{L}^{(M)} = \prod_{i=M}^{\infty} V_i.$$

For any $M \in \mathbb{N}$, the subspace

$$\mathscr{H}^M := \bigoplus_{k=0}^{M} \mathscr{H}_{(k)} \subset \mathscr{H}$$

is a counital, cosubalgebra of $(\mathscr{H}, \Delta, \varepsilon)$ and the canonical projection $\pi_M : \mathscr{H} \to \mathscr{H}^M$ is a coalgebra epimorphism.

Then the corresponding dual algebra $\big((\mathscr{H}^M)^*, \boxtimes, \varepsilon\big)$ has associated to it the truncated Lie algebra

$$\mathfrak{L}^M = \mathfrak{L}/\mathfrak{L}^{(M)} \cong \bigoplus_{i=1}^{M} V_i$$

which is a step-M nilpotent Lie algebra and its associated Lie group G^M.

C.1.1 Carnot-Carathéodory Metric

We define the dilation $(\delta_t)_{r>0}$ on the Lie algebra $\mathfrak{L}^N$ to be a collection of automorphisms of $\mathfrak{L}$ such that $\delta_s \delta_t = \delta_{st}$ and

$$\delta_t[\mathfrak{l}] = \delta_t[\mathfrak{l}_1 + ... + \mathfrak{l}_M] = t\mathfrak{l}_1 + ... + t^M \mathfrak{l}_M$$

where $\mathfrak{l}_i \in V_i$. The dilation can also be extended to the Lie group by

$$\delta_t[g] := \delta_t[\log_\boxtimes(g)].$$

A homogeneous group is any Lie group whose Lie algebra is endowed with a family of dilations.

Definition C.1.2. *A homogeneous norm on a homogeneous group G is a continuous function $\|\cdot\| : G \to \mathbb{R}^+$ such that $\|g\| = 0 \iff g = \mathbf{1}$ and $\|\delta_t[g]\| = |t| \cdot \|g\|$. A homogeneous norm is called subadditive if $\|g_1 \boxtimes g_2\| \leq \|g_1\| + \|g_2\|$ and called symmetric if $\|g^{-1}\| = \|g\|$.*

When a homogeneous norm is subadditive and symmetric, it induces a left invariant metric on G. This is traditionally called the Carnot-Carathéodory metric which we denote by d_{cc}. Finally, all homogeneous norms on a homogeneous group are equivalent.

Examples of a homogeneous norm include

$$\|g\|_{G^M} = \sum_{\tau \in \mathscr{T}_M} |\langle \log_\boxtimes(g), \tau \rangle|^{1/|\tau|} \qquad \text{and}$$
$$\|g\|_{G^M} = \sup_{\tau \in \mathscr{T}_M} |\langle \log_\boxtimes(g), \tau \rangle|^{1/|\tau|} \tag{C.1.1}$$

where $\mathscr{T}_M$ is a basis of the vector space $\mathscr{H}^M$.

Remark C.1.3. *The choice of $\mathscr{H}$ determines the structure of the rough paths that we study. When $\mathscr{H}$ is the Shuffle Hopf algebra, paths over the characters of $\mathscr{H}$ are geometric rough paths. By contrast, when $\mathscr{H}$ is the Butcher-Connes-Kreimer Hopf algebra of decorated non-planar trees, paths over the charaters of $\mathscr{H}$ are branched rough paths, see* [Gub10].

C.2 Geometric rough paths

Let $\mathscr{H}(\mathbb{R}^d)$ be the linear span of the free monoid generated by the alphabet $\{1, ..., d\}$. Let the product on $\mathscr{H}(\mathbb{R}^d)$ be the Shuffle product (denoted $\sqcup\!\sqcup$) and the coproduct on $\mathscr{H}$ be the Deconcatenation coproduct. Let $\mathcal{A}_M$ be the collection of words of length at most M, a basis for $\mathscr{H}^M(\mathbb{R}^d)$.

Definition C.2.1. *For a path $x \in C^{1-var}([0,T];\mathbb{R}^d)$, the iterated integrals of x are canonically defined using Young integration. The collection of iterated integrals of the path x is called the truncated Signature of x and is defined as*

$$S(x)_{s,t} := \mathbf{1} + \sum_{n=1}^{\infty} \int_{s\le u_1 \le ... \le u_n \le t} dx_{u_1} \otimes ... \otimes dx_{u_n} \in \mathscr{H}(\mathbb{R}^d).$$

In the same way, the truncated Signature defined by its increments

$$S_M(x)_{s,t} := \mathbf{1} + \sum_{n=1}^{M} \int_{s\le u_1 \le ... \le u_n \le t} dx_{u_1} \otimes ... \otimes dx_{u_n} \in \mathscr{H}(\mathbb{R}^d).$$

It is well known that $S_M(x)$ takes values in $G^M(V)$.

Definition C.2.2. *For $\alpha \in (0,1)$ and let M be the largest integer such that $M\alpha < 1$. A path $\mathbf{X} : [0,T] \to G^M(\mathbb{R}^d)$ is called an α-Hölder continuous geometric rough paths if*

$$\langle \mathbf{X}_{s,t}, e_A \rangle \langle \mathbf{X}_{s,t}, e_B \rangle = \langle \mathbf{X}_{s,t}, e_A \sqcup\!\sqcup e_B \rangle, \quad \langle \mathbf{X}_{s,u}, e_A \rangle = \langle \mathbf{X}_{s,t} \boxtimes \mathbf{X}_{t,u}, \Delta[e_A] \rangle$$
$$\text{and} \quad \sup_{A \in \mathcal{A}_M} \sup_{s,t\in[0,T]} \frac{\langle \mathbf{X}_{s,t}, e_A \rangle}{|t-s|^{\alpha|A|}} < \infty. \tag{C.2.1}$$

Definition C.2.3. *Denote $p = \frac{1}{\alpha}$. We define the α-Hölder rough path metric*

$$d_\alpha(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X}^{-1} \boxtimes \mathbf{Y}\|_\alpha = \sup_{s,t\in[0,T]} \frac{\left\| \mathbf{X}_{s,t}^{-1} \boxtimes \mathbf{Y}_{s,t} \right\|_{cc}}{|t-s|^\alpha}. \tag{C.2.2}$$

By quotienting with respect to $\mathbf{X}_0$, one can make this a norm. We use the convention that $\|\mathbf{X}\|_{p-var;[0,T]} = \|\mathbf{1}^{-1} \boxtimes \mathbf{X}\|_{p-var;[0,T]}$ and $\|\mathbf{X}\|_\alpha = \|\mathbf{1}^{-1} \boxtimes \mathbf{X}\|_\alpha$. We denote the metric space of α-Hölder continuous geometric rough paths to be $G\Omega_\alpha(\mathbb{R}^d)$.

Similarly, we define the homogeneous p-variation metric d_{p-var} by

$$d_{p-var;[0,T]}(\mathbf{X}, \mathbf{Y}) := \|\mathbf{X}^{-1} \boxtimes \mathbf{Y}\|_{p-var;[0,T]} := \left(\sup_{D=(t_i)} \sum_{i:t_i\in D} \left\| \mathbf{X}_{t_i,t_{i+1}}^{-1} \boxtimes \mathbf{Y}_{t_i,t_{i+1}} \right\|_{cc}^p \right)^{\frac{1}{p}}. \tag{C.2.3}$$

Remark C.2.4. *It is well known that $G\Omega_\alpha(\mathbb{R}^d)$ is the closure of the set*

$$\left\{ S_M(x) : x \in C^{1-var}([0,T];\mathbb{R}^d) \right\},$$

with respect to the α-Hölder rough path metric (C.2.2).

When studying rough paths, one can either work with p-variation or α-Hölder norms. For the most part, authors choose one and stick with it for the entirety of their work. While p-variation is slightly more general, α-Hölder allows for a wavelet representation in the Banach space which is more favourable for this work.

It is important to understand that for this paper, we work with both norms. The Hölder norm, being more restrictive, is assumed to be the bound on regularity. However, we are required to work with the p-variation in order to establish an integrability condition.

Definition C.2.5. *Let $\Delta_T = \{(s,t) : 0 \leq s \leq t \leq T\}$ denote the two-dimensional simplex. The map $\omega : \Delta_T \to \mathbb{R}^+$ is a Control if it is a continuous, non negative, super-additive function which vanishes on the diagonal.*

Example C.2.6. *Suppose that $\mathbf{X}$ is a geometric rough path with finite p-variation, so that Equation* (C.2.3) *is finite. Then $\omega_{\mathbf{X},p}(s,t) := \|\mathbf{X}\|^p_{p-var;[s,t]}$ is a control.*

The Carnot-Carathéodory metric as already described takes its structure from the Group $G^M(\mathbb{R}^{d'})$ and so is homogeneous with respect to the group dilation δ_λ. However, there is another metric that takes its structure from the vector space $T^M(\mathbb{R}^{d'})$.

For two elements $g_1, g_2 \in T^M(\mathbb{R}^{d'})$ and $i \in \{1, ..., M\}$ we have the collection of pseudo-metrics

$$\rho_i(g_1, g_2) = \sum_{\substack{A \in \mathcal{A}_M \\ |A|=i}} \Big| \langle g_1, e_A \rangle - \langle g_2, e_A \rangle \Big|. \tag{C.2.4}$$

We also have the inhomogeneous Tensor metric

$$\rho(g_1, g_2) = \max_{i=1,\dots,M} \rho_i(g_1, g_2).$$

Definition C.2.7. *Let $p = \frac{1}{\alpha} > 2$. For a fixed control ω, we define the inhomogeneous ω-modulus metric to be*

$$\rho_{p-\omega;[0,T]}(\mathbf{X}, \mathbf{Y}) := |\mathbf{X}_0 - \mathbf{Y}_0|_{T^{\lfloor p \rfloor}(\mathbb{R}^{d'})} + \max_{i=1,\dots,\lfloor p \rfloor} \sup_{s,t \in [0,T]} \frac{\rho_i(\mathbf{X}_{s,t}, \mathbf{Y}_{s,t})}{\omega(s,t)^{i/p}}. \tag{C.2.5}$$

When we additionally have that $\omega(s,t) \leq C|t-s|$ where C is a constant independent of s, t, we also have the inhomogeneous α-Hölder metric to be

$$\rho_{\alpha-H\ddot{o}l;[0,T]}(\mathbf{X}, \mathbf{Y}) := |\mathbf{X}_0 - \mathbf{Y}_0|_{T^{\lfloor p \rfloor}(\mathbb{R}^{d'})} + \max_{i=1,\dots,\lfloor p \rfloor} \sup_{s,t \in [0,T]} \frac{\rho_i(\mathbf{X}_{s,t}, \mathbf{Y}_{s,t})}{|t-s|^{\alpha i}}. \tag{C.2.6}$$

The inhomogeneous rough path metrics satisfy the simple relation

$$\rho_{p-var;[0,T]}(\mathbf{X}, \mathbf{Y}) \leq \Big(1 \vee \max_{i=1,\dots,\lfloor p \rfloor} \omega(0,T)^{i/p}\Big) \rho_{p-\omega;[0,T]}(\mathbf{X}, \mathbf{Y}) \tag{C.2.7}$$

by simple manipulation of the standard relation between p-variation and $\frac{1}{p}$-Hölder regularity, see [FV10b].

Definition C.2.8. *Let E and F be normed spaces. A map $f : E \to F$ is called γ-Lipschitz (in the sense of Stein) if f is $\lfloor \gamma \rfloor$ continuously differentiable (in the sense of Fréchet) and such that there exists a constant $M < \infty$ such that the supremum norm of the k^{th} derivative for $k = 1, ..., \lfloor \gamma \rfloor$ and the $\{\gamma\}$-Hölder norm of its $\lfloor \gamma \rfloor^{th}$ derivative are bounded by M. The smallest $M \geq 0$ satisfying this condition is the γ-Lipschitz norm of f, denoted $\|f\|_{\mathrm{Lip}^\gamma}$. The space of all such functions is denoted* $\mathrm{Lip}^\gamma(E, F)$.

We also emphasise the distinction between $\mathrm{Lip}^1_*(E, F)$, the space of functions $f : E \to F$ that are Lipschitz.

Theorem C.2.9 ([LV07])**.** *Let $V = \bigoplus V^j$ be a vector space.*

Let $\alpha < 1/2$ such that $\frac{1}{\alpha} \notin \mathbb{N}$ and $M = \lfloor \frac{1}{\alpha} \rfloor$. Suppose that $\mathbf{X}^j_t$ are α-Hölder continuous paths taking values in $G^M(V^j)$. Then $\bigoplus_j \mathbf{X}^j_t$ can be thought of as an α-Hölder continuous path taking values in $\bigoplus_j G^M(V^j)$ and there exists an extension $\mathbf{X}_t$ taking values in $G^M(V)$ that is α-Hölder continuous with respect to the Carnot norm on $G^M(V)$.

C.2.1 Translation of rough paths

We define the map $\#: G^M(\mathbb{R}^d \oplus \mathbb{R}^d) \to G^M(\mathbb{R}^d)$ to be the unique homomorphism such that for $v_1, v_2 \in \mathbb{R}^d$, $\#[\exp_{\boxtimes}(v_1 \oplus v_2)] = \exp_{\boxtimes}(v_1 + v_2)$.

Definition C.2.10. *Let $\alpha, \beta > 0$ such that $\alpha + \beta > 1$. Let M be the greatest integer such that $M\alpha < 1$. Let $(\mathbf{X}, h) \in C^\alpha([0,T]; G^M(\mathbb{R}^d)) \times C^\beta([0,T]; \mathbb{R}^d)$. We define the Translation of the rough path $\mathbf{X}$ by the path h, denoted $T^h(\mathbf{X}) \in C^\alpha([0,T]; G^M(\mathbb{R}^d))$ to be*

$$T^h(\mathbf{X}) = \#\Big[S_M(\mathbf{X} \oplus h)\Big]$$

Lemma C.2.11. *Let $\alpha \in (\frac{1}{4}, \frac{1}{2}]$ and let $\mathbf{X} \in G\Omega_\alpha(\mathbb{R}^d)$. Let $p = \frac{1}{\alpha}$ and let $1 = \frac{1}{p} + \frac{1}{q}$. Let $h : [0,T] \to \mathbb{R}^d$ satisfy that*

$$\|h\|_{q,\alpha,[0,T]} := \sup_{t,s \in [0,T]} \frac{\|h\|_{q-var,[s,t]}}{|t-s|^\alpha} < \infty.$$

Then:

1. *There exists $C = C(p,q) > 0$ such that*

$$\|T^h(\mathbf{X})\|_\alpha \leq C\Big(\|\mathbf{X}\|_\alpha + \|h\|_{q,\alpha,[0,T]}\Big) \tag{C.2.8}$$

2. *The homogeneous rough path metric d_α is T^h-invariant.*

Proof. A proof of Equation (C.2.8) can be found in [CLL13].

Using that $\#$ is a Group homomorphism, we have

$$\begin{aligned}
\Big(T^h(\mathbf{X})_{s,t}\Big)^{-1} \boxtimes T^h(\mathbf{Y})_{s,t} =& \Big(\#\Big[S_M(\mathbf{X} \oplus h)\Big]\Big)^{-1}_{s,t} \boxtimes \#\Big[S_M(\mathbf{Y} \oplus h)\Big] \\
=& \#\Big[S_M(\mathbf{X} \oplus h)^{-1}_{s,t} \boxtimes S_M(\mathbf{Y} \oplus h)_{s,t}\Big] \\
=& \#\Big[S_M\Big(\mathbf{X}^{-1} \boxtimes \mathbf{Y} \oplus 0\Big)_{s,t}\Big] = \mathbf{X}^{-1}_{s,t} \boxtimes \mathbf{Y}_{s,t}
\end{aligned}$$

□

C.2.2 The lift of a Gaussian Process

Gaussian processes have a natural lift for their signiture. It is shown in [FH14] that one can solve the iterated integral of a Gaussian process by approximating the process pathwise and showing that the approximation converges in mean square and almost surely. In particular, the iterated integral of a Gaussian process is an element on the second Wiener-Itô chaos expansion.

The key to this result is the regularity of the covariance function of the Gaussian process. Provided the covariance function is adequately continuous, the existence of the lift to the signiture is assured.

In [FV10a], the authors prove that when the covariance operator of the Gaussian satisfies a p-variation condition, the path of the Gaussian can be lifted to a rough path with p-variation and α-Hölder continuity in the rough path sense.

Assumption C.2.12. *Let $\mathcal{L}^W$ be the law of a d-dimensional, continuous centred Gaussian process with independent components and covariance covariance operator $\mathcal{R}$ such that $\exists \varrho \in [1, 3/2)$ and $M < \infty$ with*

$$\|\mathcal{R}\|_{\varrho;[s,t]^2} \leq M|t-s|^{1/\varrho}.$$

C.3 Controlled rough paths

A controlled rough path, first introduced in [Gub04], provides a path that is known to be adequately regular enough to be integrable with respect to a rough path.

Let V and U be vector spaces and denote by $L(V,U)$ the space of Linear operators from V to U. We define $T(V^*,U) := \bigoplus_{n=0}^{\infty} L\big((V^*)^{\otimes n}, U\big)$ and use the convention that $L\big((V^*)^{\otimes 0}, U\big) = U$. As earlier, we are interested in the case where $V = \mathbb{R}^{d'}$ and $U = \mathbb{R}^d$.

Given an element $\mathbf{X} \in T(V^*)$ and $\mathbb{Y} \in T(V^*,U)$, we naturally obtain $\mathbb{Y}\mathbf{X} \in U$. Also, in practice we work in the truncated tensor algebra $T^M(V^*,U) := \bigoplus_{n=0}^{M} L\big((V^*)^{\otimes n}, U\big)$ obtained by quotienting with respect to the ideal $\bigoplus_{n=M+1}^{\infty} L\big((V^*)^{\otimes n}, U\big)$.

Definition C.3.1. *Let $\alpha \in (0,1/2)$, let M be the smallest integer such that $M\alpha < 1$ and let $\mathbf{X} \in G\Omega_\alpha(V)$. Let $\mathcal{A}_v$ be the alphabet of V.*

A $\mathbf{X}$-controlled rough path $\mathbb{Y} : [0,T] \to T^{M-1}(V,U)$ and a remainder term $R : \Delta_T \to T^{M-1}(V,U)$ is any path such that for any word A of the alphabet for $\mathcal{A}_v$

$$\langle \mathbb{Y}_t, e_A\rangle - \langle \mathbb{Y}_s, \mathbf{X}_{s,t} \boxtimes e_A\rangle = \langle R_{s,t}, e_A\rangle,$$

where

$$\sup_{t,s\in[0,T]} \frac{|\langle R_{s,t}, e_A\rangle|}{|t-s|^{(M-|A|)\alpha}} < \infty.$$

The space of $\mathbf{X}$-controlled rough paths, denoted $\mathcal{D}_{\mathbf{X}}^{M\alpha}([0,T];U)$ is the vector space of all $\mathbf{X}$-Controlled paths with the norm

$$\|\mathbb{Y}\|_{\mathbf{X},M\alpha} = \sum_{A\in\mathcal{A}_M\setminus\{\varepsilon\}} \Big\|\langle \mathbb{Y}, e_A\rangle\Big\|_{|A|\alpha-H\ddot{o}l;[0,T]} + \Big\|\langle R, e_\varepsilon\rangle\Big\|_{M\alpha-H\ddot{o}l;[0,T]}.$$

Given an $\mathbf{X}$-controlled rough path $\mathbb{Y}$ taking values on $L(V,U)$, we define the integral

$$\int_0^T Y_t d\mathbf{X}_t = \lim_{|D|\to 0} \sum_{i:t_i\in D} \Big\langle \mathbb{Y}_{t_i}, \mathbf{X}_{t_i,t_{i+1}} \Big\rangle_{T^{M-1}\big(V,L(V,U)\big),T^M\big(V\big)}$$

taking values in U.

Definition C.3.2. *The Shuffle product over $T(V)$ can be represented as two Left and Right Half-shuffle products $e_A \sqcup\!\sqcup\, e_B = e_A \prec e_B + e_A \succ e_B$ that satisfy the identities*

$$\begin{aligned}
(e_A \prec e_B) \prec e_C &= e_A \prec (e_B \sqcup\!\sqcup\, e_C),\\
(e_A \succ e_B) \prec e_C &= e_A \succ (e_B \prec e_C),\\
(e_A \sqcup\!\sqcup\, e_B) \succ e_C &= e_A \succ (e_B \succ e_C).
\end{aligned}$$

Using the additional identity $e_A \prec e_B = e_B \succ e_A$, equivalent to commutivity of $\sqcup\!\sqcup$, we observe that the Left and Right Half-shuffles satisfy a Left and Right Zinbiel identity. Thus $\succ$ and $\prec$ are sometimes referred to as Paraproducts. For any geometric rough path and any two words A and B we have

$$\int_s^t \langle \mathbf{X}_{s,r}, e_A\rangle d\langle \mathbf{X}_{s,r}, e_B\rangle = \langle \mathbf{X}_{s,t}, e_A \succ e_B\rangle = \langle \delta_\succ[\mathbf{X}_{s,t}], e_A \otimes e_B\rangle. \tag{C.3.1}$$

where $\delta_\succ$ is the Right Half-Unshuffle. Using the Right Half-Unshuffle, we are able to "stitch" two controlled rough paths together to obtain an object that will satisfy the Sewing Lemma, providing us with a meaningful way to integrate a controlled rough path with respect to another controlled rough path.

Theorem C.3.3. *Let $\mathbb{Y}$ and $\mathbb{Z}$ be $\mathbf{X}$-controlled rough paths. Then by exploiting Equation* (C.3.1) *we obtain*

$$\begin{aligned}
&\int_0^T Y_t \otimes dZ_t\\
&= \lim_{|D|\to 0} \sum_{i:t_i\in D} Y_{t_i} \otimes Z_{t_i,t_{i+1}} + \Big\langle \Big(\mathbb{Y}_{t_i} - Y_{t_i}\Big) \otimes \Big(\mathbb{Z}_{t_i} - Z_{t_i}\Big), \delta_\succ[\mathbf{X}_{t_i,t_{i+1}}]\Big\rangle.
\end{aligned}$$

In a similar fashion, we obtain

$$\begin{aligned}&\int_s^t Y_{s,r}\otimes dZ_r\\&=\lim_{|D|\to 0}\sum_{i:t_i\in D}\Big\langle\Big(\mathbb{Y}_{t_i}-Y_{t_i}\Big)\otimes\Big(\mathbb{Z}_{t_i}-Z_{t_i}\Big),\delta_{\succ}[\mathbf{X}_{t_i,t_{i+1}}]\Big\rangle.\end{aligned}$$

Given an $\mathbf{X}$*-controlled rough path* $\mathbb{Y}$*, one can extend it to a rough path* $\mathbf{Y}$ *taking values in* $G^M(U)$*. Define the path* $\mathbf{Y}:[0,T]\to G^M(U)$ *by*

$$\mathbf{Y}_{s,t}=\mathbf{1}+\sum_{k=1}^{M}\lim_{|D|\to 0}\sum_{i:t_i\in D}\Big\langle\big(\mathbb{Y}_{t_i}-Y_{t_i}\big)^{\otimes k},(\delta_{\succ})^k[\mathbf{X}_{t_i,t_{i+1}}]\Big\rangle \tag{C.3.2}$$

where the iterated coproduct $(\delta_{\succ})^k:T^M(V^*)\to T^M(V^*)^{\otimes k}$ *is defined inductively by*

$$(\delta_{\succ})^2=((\delta_{\succ})\otimes I)\delta_{\succ},\quad(\delta_{\succ})^{k+1}=((\delta_{\succ})\otimes I^{\otimes k})\delta_{\succ}^k.$$

Proof of Theorem C.3.3. The ideas behind this proof are well understood, although to the best of the authors knowledge have not been written using the language of Zinbiel algebras before.

Firstly,

$$\int_s^t Y_r\otimes dZ_r=Y_s\otimes Z_{s,t}+\int_s^t Y_{s,r}\otimes dZ_r$$

and from the definition of controlled rough paths we have

$$\begin{aligned}Y_s\otimes Z_{s,t}=&Y_s\otimes\langle\mathbb{Z}_s,\mathbf{X}_{s,t}-\mathbf{1}\rangle+Y_s\otimes\langle R^Z_{s,t},e_\varepsilon\rangle,\\Y_{s,r}=&\langle\mathbb{Y}_s,\mathbf{X}_{s,r}-\mathbf{1}\rangle+\langle R^Y_{s,r},e_\varepsilon\rangle.\end{aligned}$$

Thus

$$\begin{aligned}&\int_s^t Y_{s,r}\otimes dZ_r\\&=\left(\Big\langle\mathbb{Y}_s-Y_s,\int_s^t\mathbf{X}_{s,r}\Big\rangle+\int_s^t\langle R^Y_{s,r},e_\varepsilon\rangle\right)\otimes\left(\Big\langle\mathbb{Z}_s-Z_s,d\mathbf{X}_{s,r}\Big\rangle+\langle dR^Z_{s,r},e_\varepsilon\rangle\right),\\&=\left\langle\Big(\mathbb{Y}_s-Y_s\Big)\otimes\Big(\mathbb{Z}_s-Z_s\Big),\int_s^t\mathbf{X}_{s,r}d\mathbf{X}_{s,r}\right\rangle+o\Big(|t-s|\Big)\end{aligned}$$

as $|t-s|\to 0$ where we use the identity from Equation (C.3.1) and the regularity of Definition C.3.1. Similarly

$$\int_s^t Y_r\otimes dZ_r=Y_s\otimes Z_{s,t}+\Big\langle(\mathbb{Y}_s-Y_s)\otimes(\mathbb{Z}_s-Z_s),\delta_{\succ}[\mathbf{X}_{s,t}]\Big\rangle+o(|t-s|).$$

Motivated by this, we verify the conditions of the Sewing Lemma (see [FH14, Lemma 4.2]) with

$$\Xi_{s,t}:=Y_s\otimes Z_{s,t}+\Big\langle(\mathbb{Y}_s-Y_s)\otimes(\mathbb{Z}_s-Z_s),\delta_{\succ}[\mathbf{X}_{s,t}]\Big\rangle.$$

Thus for $s<t<u\in[0,T]$,

$$\begin{aligned}\delta\Xi_{s,t,u}=&\Xi_{s,u}-\Xi_{s,t}-\Xi_{t,u}\\=&-Y_{s,t}\otimes Z_{t,u}+\left\langle\Big((\mathbb{Y}-Y)\otimes(\mathbb{Z}-Z)\Big)_{s,t},\delta_{\succ}[\mathbf{X}_{t,u}]\right\rangle\\&+\sum_{A,B}\Big((\mathbb{Y}_s-Y_s)\otimes(\mathbb{Z}_s-Z_s)\Big)[e_A\otimes e_B]\Big\langle\mathbf{X}_{s,t}\otimes\mathbf{X}_{t,u},\overline{\Delta}[e_A\succ e_B]\Big\rangle\end{aligned}$$

where $\overline{\Delta}$ is the reduced Coproduct. Next, we substitute in for the increments using the identities

$$\langle \mathbb{Y}_t, e_A \rangle - \langle \mathbb{Y}_s, \mathbf{X}_{s,t} \boxtimes e_A \rangle = \langle R_{s,t}, e_A \rangle,$$
$$\langle \mathbb{Z}_t, e_B \rangle - \langle \mathbb{Z}_s, \mathbf{X}_{s,t} \boxtimes e_B \rangle = \langle R_{s,t}, e_B \rangle.$$

Next, we use Sweedler notation to represent the identity

$$\overline{\Delta}[e_A \succ e_B] = \sum_{A',A''} \sum_{B',B''} = e_{A' \sqcup B'} \otimes e_{A'' \succ B''}.$$

Therefore

$$\sup_{s,t,u \in [0,T]} \frac{\delta \Xi_{s,t,u}}{|u-s|^{M\alpha}} = o\Big(|u-s| \Big).$$

The ideas behind this proof are well understood (see [LCL07, p.74]) where Y is the solution to a linear rough differential equation, although to the best of the authors' knowledge they have not been written before using the language of Zinbiel algebras and for general controlled rough paths. We refer the reader to the forthcoming preprint [CDFL20], where a proof is given of this result. □

Bibliography

[ADP90] J.-P. Aubin and G. Da Prato, *Stochastic viability and invariance*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **17** (1990), no. 4, 595–613. MR1093711

[AILVZ08] F. Aurzada, I. A. Ibragimov, M. A. Lifshits, and Kh. Ya. Van Zanten, *Small deviations of smooth stationary Gaussian processes*, Teor. Veroyatn. Primen. **53** (2008), no. 4, 788–798. MR2766145

[ALL09] F. Aurzada, M. Lifshits, and W. Linde, *Small deviations of stable processes and entropy of the associated random operators*, Bernoulli **15** (2009), no. 4, 1305–1334. MR2597594

[Aur07] F. Aurzada, *Metric entropy and the small deviation problem for stable processes*, Probab. Math. Statist. **27** (2007), no. 2, 261–274. MR2445997

[BAGL94] G. Ben Arous, M. Grădinaru, and M. Ledoux, *Hölder norms and the support theorem for diffusions*, Ann. Inst. H. Poincaré Probab. Statist. **30** (1994), no. 3, 415–436. MR1288358

[Bal86] P. Baldi, *Large deviations and functional iterated logarithm law for diffusion processes*, Probab. Theory Relat. Fields **71** (1986), no. 3, 435–453. MR824713

[BAL94] G. Ben Arous and M. Ledoux, *Grandes déviations de Freidlin-Wentzell en norme hölderienne*, Séminaire de Probabilités, XXVIII, 1994, pp. 293–299. MR1329119

[BC19] X. Bay and J.-C. Croix, *Karhunen-Loève decomposition of Gaussian measures on Banach spaces*, Probability and Mathematical Statistics **39** (2019), no. 2.

[BCCP98] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti, *A non-Maxwellian steady distribution for one-dimensional granular media*, J. Statist. Phys. **91** (1998), no. 5-6, 979–990. MR1637274

[BCD18] I. Bailleul, R. Catellier, and F. Delarue, *Mean field rough differential equations*, ArXiv e-prints (February 2018), available at `1802.05882`.

[BCD19] I. Bailleul, R. Catellier, and F. Delarue, *Propagation of Chaos for Mean field rough differential equations*, ArXiv e-prints (July 2019), available at `1907.00578`.

[BCD20] I. Bailleul, R. Catellier, and F. Delarue, *Solving mean field rough differential equations*, Electronic Journal of Probability **25** (2020).

[BCM07] M. Burger, V. Capasso, and D. Morale, *On an aggregation model with long and short range interactions*, Nonlinear Anal. Real World Appl. **8** (2007), no. 3, 939–958. MR2307761

[BDF12] A. Budhiraja, P. Dupuis, and M. Fischer, *Large deviation properties of weakly interacting processes via weak convergence methods*, Ann. Probab. **40** (2012), no. 1, 74–102. MR2917767

[BGG13] F. Bolley, I. Gentil, and A. Guillin, *Uniform convergence to equilibrium for granular media*, Arch. Ration. Mech. Anal. **208** (2013), no. 2, 429–445. MR3035983

[Bit13] J. Bittencourt, *Fundamentals of plasma physics*, Springer Science & Business Media, 2013.

[BJT11] M. Bossy, J.-F. Jabir, and D. Talay, *On conditional McKean Lagrangian stochastic models*, Probab. Theory Related Fields **151** (2011), no. 1-2, 319–351. MR2834721

[BL16] J. M. Borwein and S. B. Lindstrom, *Meetings with Lambert $\mathcal{W}$ and other special functions in optimization and analysis*, Pure Appl. Funct. Anal. **1** (2016), no. 3, 361–396. MR3622396

[BLG14] E. Boissard and T. Le Gouic, *On the mean speed of convergence of empirical and occupation measures in Wasserstein distance*, 2014, pp. 539–563. MR3189084

[Bog98] V. I. Bogachev, *Gaussian measures*, Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998. MR1642391

[Boi11] E. Boissard, *Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance*, Electron. J. Probab. **16** (2011), no. 83, 2296–2333. MR2861675

[Bol08] F. Bolley, *Separability and completeness for the Wasserstein distance*, Lecture Notes in Math., vol. 1934, Springer, Berlin, 2008. MR2483740

[BQRT10] R. Buckdahn, M. Quincampoix, C. Rainer, and J. Teichmann, *Another proof for the equivalence between invariance of closed sets with respect to stochastic and deterministic systems*, Bull. Sci. Math. **134** (2010), no. 2, 207–214. MR2592970

[BR92] P. Baldi and B. Roynette, *Some exact equivalents for the Brownian motion in Hölder norm*, Probab. Theory Related Fields **93** (1992), no. 4, 457–484. MR1183887

[BRTV98] S. Benachour, B. Roynette, D. Talay, and P. Vallois, *Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos*, Stochastic Process. Appl. **75** (1998), no. 2, 173–201. MR1632193

[BT11] J. Binney and S. Tremaine, *Galactic dynamics*, Vol. 20, Princeton university press, 2011.

[Car16] R. Carmona, *Lectures of BSDEs, stochastic control, and stochastic differential games with financial applications*, SIAM, 2016.

[CD17a] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications I*, 1st ed., Probability Theory and Stochastic Modelling, vol. 84, Springer International Publishing, 2017.

[CD17b] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications II*, 1st ed., Probability Theory and Stochastic Modelling, vol. 84, Springer International Publishing, 2017.

[CDFL20] T. Cass, B. Driver, E. Ferrucci, and C. Litterer, *Intrinsic rough paths on banach manifolds*, internal preprint (2020).

[CDFM18] M. Coghi, J.-D. Deuschel, P. Friz, and M. Maurelli, *Pathwise McKean-Vlasov theory with additive noise*, arXiv preprint arXiv:1812.11773 (2018).

[CDL13] R. Carmona, F. Delarue, and A. Lachapelle, *Control of McKean-Vlasov dynamics versus mean field games*, Math. Financ. Econ. **7** (2013), no. 2, 131–166. MR3045029

[Cer01] S. Cerrai, *Second order PDE's in finite and infinite dimension*, Lecture Notes in Mathematics, vol. 1762, Springer-Verlag, Berlin, 2001. A probabilistic approach. MR1840644

[CF10] T. Cass and P. Friz, *Densities for rough differential equations under Hörmander's condition*, Ann. of Math. (2) **171** (2010), no. 3, 2115–2141. MR2680405

[CF18] K. Chouk and P. K. Friz, *Support theorem for a singular SPDE: the case of gPAM*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 1, 202–219. MR3765886

[CFN97] M. E. Caballero, B. Fernández, and D. Nualart, *Composition of skeletons and support theorems*, Stochastic differential and difference equations (Györ, 1996), 1997, pp. 21–33. MR1636824

[CGM08] P. Cattiaux, A. Guillin, and F. Malrieu, *Probabilistic approach for granular media equations in the non-uniformly convex case*, Probab. Theory Related Fields **140** (2008), no. 1-2, 19–40. MR2357669

[CHLT15] T. Cass, M. Hairer, C. Litterer, and S. Tindel, *Smoothness of the density for solutions to Gaussian rough differential equations*, Ann. Probab. **43** (2015), no. 1, 188–239. MR3298472

[CK19] R. Cont and A. Kalinin, *On the support of solutions to stochastic differential equations with path-dependent coefficients*, Stochastic processes and their applications (2019). In Press.

[CL15] T. Cass and T. Lyons, *Evolving communities with individual preferences*, Proc. Lond. Math. Soc. (3) **110** (2015), no. 1, 83–107. MR3299600

[CLL13] T. Cass, C. Litterer, and T. Lyons, *Integrability and tail estimates for Gaussian rough differential equations*, Ann. Probab. **41** (2013), no. 4, 3026–3050. MR3112937

[CM18] D. Crisan and E. McMurray, *Smoothing properties of McKean-Vlasov SDEs*, Probab. Theory Related Fields **171** (2018), no. 1-2, 97–148. MR3800831

[CN19] M. Coghi and T. Nilssen, *Rough nonlocal diffusions*, arXiv preprint arXiv:1905.07270 (2019).

[CRS19] T. Cass, G. d. Reis, and W. Salkeld, *The support of mckean vlasov equations driven by brownian motion*, arXiv preprint arXiv:1911.01992 (2019).

[CS19] J. G. Conlon and A. Schlichting, *A non-local problem for the Fokker-Planck equation related to the Becker-Döring model*, Discrete Contin. Dyn. Syst. **39** (2019), no. 4, 1821–1889. MR3927496

[CS90] B. Carl and I. Stephani, *Entropy, compactness and the approximation of operators*, Cambridge University Press, 1990.

[DE97] P. Dupuis and R. S. Ellis, *A weak convergence approach to the theory of large deviations*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1997. A Wiley-Interscience Publication. MR1431744

[DFMS03] S. Dereich, F. Fehringer, A. Matoussi, and M. Scheutzow, *On the link between small ball probabilities and the quantization problem for Gaussian measures on Banach spaces*, J. Theoret. Probab. **16** (2003), no. 1, 249–265. MR1956830

[DFMS18] J.-D. Deuschel, P. K. Friz, M. Maurelli, and M. Slowik, *The enhanced Sanov theorem and propagation of chaos*, Stochastic Process. Appl. **128** (2018), no. 7, 2228–2269. MR3804792

[DG87] D. A. Dawson and J. Gärtner, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics **20** (1987), no. 4, 247–308. MR885876

[DIRT15] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré, *Global solvability of a networked integrate-and-fire model of McKean-Vlasov type*, Ann. Appl. Probab. **25** (2015), no. 4, 2096–2133. MR3349003

[DM14] D. Dobbs and T. Melcher, *Small deviations for time-changed Brownian motions and applications to second-order chaos*, Electron. J. Probab. **19** (2014), no. 85, 23. MR3263642

[DNØkP09] G. Di Nunno, B. Ø ksendal, and F. Proske, *Malliavin calculus for Lévy processes with applications to finance*, Universitext, Springer-Verlag, Berlin, 2009. MR2460554

[DP06] G. Da Prato, *An introduction to infinite-dimensional analysis*, Springer Science & Business Media, 2006.

[DPdH96] P. Dai Pra and F. den Hollander, *McKean-Vlasov limit for interacting random processes in random media*, J. Statist. Phys. **84** (1996), no. 3-4, 735–772. MR1400186

[dR20] P. C. de Raynal, *Strong well posedness of mckean–vlasov stochastic differential equations with hölder drift*, Stochastic Processes and their Applications **130** (2020), no. 1, 79–107.

[dRES18] G. dos Reis, S. Engelhardt, and G. Smith, *Simulation of mckean vlasov sdes with super linear growth*, arXiv preprint arXiv:1808.05530 (2018).

[dRST18] G. dos Reis, G. Smith, and P. Tankov, *Importance sampling for mckean-vlasov sdes*, arXiv preprint arXiv:1803.09320 (2018).

[dRST19] G. dos Reis, W. Salkeld, and J. Tugaut, *Freidlin-Wentzell LDP in path space for McKean-Vlasov equations and the functional iterated logarithm law*, Ann. Appl. Probab. **29** (2019), no. 3, 1487–1540. MR3914550

[DS55] C. W. Dunnett and M. Sobel, *Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution*, Biometrika **42** (1955), 258–260. MR68169

[DS88] N. Dunford and J. T. Schwartz, *Linear operators. Part III*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1988. Spectral operators, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1971 original, A Wiley-Interscience Publication. MR1009164

[DZ10] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition. MR2571413

[Edd00] M. Eddahbi, *Large deviation estimates and functional law for diffusions in modulus spaces*, Random Oper. Stochastic Equations **8** (2000), no. 4, 373–396. MR1797027

[Elw92] K. D. Elworthy, *Stochastic flows on Riemannian manifolds*, Diffusion processes and related problems in analysis, volume ii: Stochastic flows, 1992, pp. 37–72.

[EN02] M. Eddahbi and M. N'zi, *Strassen's local law for diffusion processes under strong topologies*, Acta Math. Vietnam. **27** (2002), no. 2, 151–163. MR1939953

[ET96] D. E. Edmunds and H. Triebel, *Function spaces, entropy numbers, differential operators*, Cambridge Tracts in Mathematics, vol. 120, Cambridge University Press, Cambridge, 1996. MR1410258

[FH14] P. K. Friz and M. Hairer, *A course on rough paths*, Universitext, Springer, Cham, 2014. With an introduction to regularity structures. MR3289027

[FHH+01] M. Fabian, P. Habala, P. Hájek, V. Montesinos Santalucía, J. Pelant, and V. Zizler, *Functional analysis and infinite-dimensional geometry*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 8, Springer-Verlag, New York, 2001. MR1831176

[FJ+17] N. Fournier, B. Jourdain, et al., *Stochastic particle approximation of the keller–segel equation and two-dimensional generalization of bessel processes*, The Annals of Applied Probability **27** (2017), no. 5, 2807–2861.

[FLLL01] E. Fournié, J.-M. Lasry, J. Lebuchoux, and P.-L. Lions, *Applications of Malliavin calculus to Monte-Carlo methods in finance. II*, Finance and Stochastics **5** (2001Apr), no. 2, 201–236.

[FLL+99] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi, *Applications of Malliavin calculus to Monte-Carlo methods in finance*, Finance and Stochastics **3** (1999Aug), no. 4, 391–412.

[FLS06] P. Friz, T. Lyons, and D. Stroock, *Lévy's area under conditioning*, Ann. Inst. H. Poincaré Probab. Statist. **42** (2006), no. 1, 89–101. MR2196973

[FR11] P. Friz and S. Riedel, *Convergence rates for the full Brownian rough paths with applications to limit theorems for stochastic flows*, Bull. Sci. Math. **135** (2011), no. 6-7, 613–628. MR2838093

[FTJ79] T. Figiel and N. Tomczak-Jaegermann, *Projections onto Hilbertian subspaces of Banach spaces*, Israel J. Math. **33** (1979), no. 2, 155–171. MR571251

[FTT14] D. Filipović, S. Tappe, and J. Teichmann, *Invariant manifolds with boundary for jump-diffusions*, Electron. J. Probab. **19** (2014), no. 51, 28. MR3227060

[FV10a] P. Friz and N. Victoir, *Differential equations driven by Gaussian signals*, 2010, pp. 369–413. MR2667703

[FV10b] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010. Theory and applications. MR2604669

[FW12] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, Third, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 260, Springer, Heidelberg, 2012. Translated from the 1979 Russian original by Joseph Szücs. MR2953753

[Gan93] N. Gantert, *An inversion of Strassen's law of the iterated logarithm for small time*, Ann. Probab. **21** (1993), no. 2, 1045–1049. MR1217579

[Gao04] F. Gao, *Entropy of absolute convex hulls in Hilbert spaces*, Bull. London Math. Soc. **36** (2004), no. 4, 460–468. MR2069008

[Gao08] F. Gao, *Entropy estimate for k-monotone functions via small ball probability of integrated Brownian motion*, Electron. Commun. Probab. **13** (2008), 121–130. MR2386068

[GEO+72] S D. Gupta, M. L Eaton, I. Olkin, M Perlman, L. J Savage, and M. Sobel, *Inequalities on the probability content of convex regions for elliptically contoured distributions*, Proceedings of the sixth berkeley symposium on mathematical statistics and probability (univ. california, berkeley, calif., 1970/1971), 1972, pp. 241–265.

[GHT03] F. Gao, J. Hannig, and F. Torcaso, *Integrated Brownian motions and exact L_2-small balls*, Ann. Probab. **31** (2003), no. 3, 1320–1337. MR1989435

[GL00] S. Graf and H. Luschgy, *Foundations of quantization for probability distributions*, Lecture Notes in Mathematics, vol. 1730, Springer-Verlag, Berlin, 2000. MR1764176

[GLP03] S. Graf, H. Luschgy, and G. Pagès, *Functional quantization and small ball probabilities for Gaussian processes*, J. Theoret. Probab. **16** (2003), no. 4, 1047–1062 (2004). MR2033197

[GNSS95] I. Gyöngy, D. Nualart, and M. Sanz-Solé, *Approximation and support theorems in modulus spaces*, Probab. Theory Related Fields **101** (1995), no. 4, 495–509. MR1327223

[GP90] I. Gyöngy and T. Pröhle, *On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem*, Comput. Math. Appl. **19** (1990), no. 1, 65–70. MR1026782

[GR08] P. Guasoni and S. Robertson, *Optimal importance sampling with explicit formulas in continuous time*, Finance Stoch. **12** (2008), no. 1, 1–19. MR2362149

[GS16] C. Geiss and A. Steinicke, *Malliavin derivative of random functions and applications to Lévy driven BSDEs*, Electron. J. Probab. **21** (2016), Paper No. 10, 28. MR3485352

[Gub04] M. Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140. MR2091358

[Gub10] M. Gubinelli, *Ramification of rough paths*, J. Differential Equations **248** (2010), no. 4, 693–721. MR2578445

[HIP08] S. Herrmann, P. Imkeller, and D. Peithmann, *Large deviations and a Kramers' type law for self-stabilizing diffusions*, Ann. Appl. Probab. **18** (2008), no. 4, 1379–1423. MR2434175

[HIPP14] S. Herrmann, P. Imkeller, I. Pavlyukevich, and D. Peithmann, *Stochastic resonance*, Mathematical Surveys and Monographs, vol. 194, American Mathematical Society, Providence, RI, 2014. A mathematical approach in the small noise limit. MR3155413

[HJK11] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, *Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **467** (2011), no. 2130, 1563–1576. MR2795791

[HK02] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence models, analysis, and simulation*, Journal of artificial societies and social simulation **5** (2002), no. 3.

[HKR19] K. Hu, A. Kazeykina, and Z. Ren, *Mean-field langevin system, optimal control and deep neural networks*, arXiv preprint arXiv:1909.07278 (2019).

[HS19] M. Hairer and P. Schönbauer, *The support of singular stochastic PDEs*, arXiv e-prints (2019Sep), arXiv:1909.05526, available at 1909.05526.

[HŠS18] W. Hammersley, D. Šiška, and L. Szpruch, *Mckean-vlasov sdes under measure dependent lyapunov conditions*, arXiv preprint arXiv:1802.03974 (2018).

[IdRS19] P. Imkeller, G. dos Reis, and W. Salkeld, *Differentiability of SDEs with drifts of super-linear growth*, Electron. J. Probab. **24** (2019), Paper No. 3, 43. MR3916323

[IMPR16] P. Imkeller, T. Mastrolia, D. Possamaï, and A. Réveillac, *A note on the Malliavin-Sobolev spaces*, Statist. Probab. Lett. **109** (2016), 45–53. MR3434959

[Kac56] M. Kac, *Foundations of kinetic theory*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, 1956, pp. 171–197. MR0084985

[Kha67] C. G. Khatri, *On certain inequalities for normal distributions and their applications to simultaneous confidence bounds*, Ann. Math. Statist. **38** (1967), 1853–1867. MR220392

[KL17] T. Kühn and W. Linde, *Gaussian approximation numbers and metric entropy*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **457** (2017), no. Veroyatnostʹi Statistika. 25, 194–210. MR3723582

[KL93a] J. Kuelbs and W. V. Li, *Metric entropy and the small ball problem for Gaussian measures*, J. Funct. Anal. **116** (1993), no. 1, 133–157. MR1237989

[KL93b] J. Kuelbs and W. V. Li, *Small ball estimates for Brownian motion and the Brownian sheet*, J. Theoret. Probab. **6** (1993), no. 3, 547–577. MR1230346

[Kle13] O. Kley, *Kuelbs-Li inequalities and metric entropy of convex hulls*, J. Theoret. Probab. **26** (2013), no. 3, 649–665. MR3090544

[KLL94] J. Kuelbs, W. V. Li, and W. Linde, *The Gaussian measure of shifted balls*, Probab. Theory Related Fields **98** (1994), no. 2, 143–162. MR1258983

[KLS95] J. Kuelbs, W. V. Li, and Q. M. Shao, *Small ball probabilities for Gaussian processes with stationary increments under Hölder norms*, J. Theoret. Probab. **8** (1995), no. 2, 361–386. MR1325856

[Øks03] B. Øksendal, *Stochastic differential equations: An introduction with applications*, Hochschultext / Universitext, Springer, 2003.

[KS71] E. Keller and L. Segel, *Model for chemotaxis*, Journal of theoretical biology **30** (1971), no. 2, 225–234.

[Kul09] A. M. Kulik, *Exponential ergodicity of the solutions to SDE's with a jump noise*, Stochastic Process. Appl. **119** (2009), no. 2, 602–632. MR2494006

[Kun90] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990. MR1070361 (91m:60107)

[Lal10] T. Laloë, *L_1-quantization and clustering in Banach spaces*, Math. Methods Statist. **19** (2010), no. 2, 136–150. MR2723146

[LaM17] R. Latał a and D. Matlak, *Royen's proof of the Gaussian correlation inequality*, Geometric aspects of functional analysis, 2017, pp. 265–275. MR3645127

[LCL07] T. J Lyons, M. Caruana, and T. Lévy, *Differential equations driven by rough paths*, Springer, 2007.

[Li99] W. V. Li, *A Gaussian correlation inequality and its applications to small ball probabilities*, Electron. Comm. Probab. **4** (1999), 111–118. MR1741737

[Lif13] M. A. Lifshits, *Gaussian random functions*, Vol. 322, Springer Science & Business Media, 2013.

[LL04] W. V. Li and W. Linde, *Small deviations of stable processes via metric entropy*, J. Theoret. Probab. **17** (2004), no. 1, 261–284. MR2054588

[LL99] W. V. Li and W. Linde, *Approximation, metric entropy and small ball estimates for Gaussian measures*, Ann. Probab. **27** (1999), no. 3, 1556–1578. MR1733160

[LP02] H. Luschgy and G. Pagès, *Functional quantization of Gaussian processes*, J. Funct. Anal. **196** (2002), no. 2, 486–531. MR1943099

[LP04] H. Luschgy and G. Pagès, *Sharp asymptotics of the Kolmogorov entropy for Gaussian measures*, J. Funct. Anal. **212** (2004), no. 1, 89–120. MR2065239

[LQ02] T. Lyons and Z. Qian, *System control and rough paths*, Oxford University Press, 2002.

[LQZ02] M. Ledoux, Z. Qian, and T. Zhang, *Large deviations and support theorem for diffusion processes via rough paths*, Stochastic Process. Appl. **102** (2002), no. 2, 265–283. MR1935127

[LS01] W. V Li and Q-M Shao, *Gaussian processes: inequalities, small ball probabilities and applications*, Handbook of Statistics **19** (2001), 533–597.

[LT13] M. Ledoux and M. Talagrand, *Probability in banach spaces: isoperimetry and processes*, Springer Science & Business Media, 2013.

[LV07] T. Lyons and N. Victoir, *An extension theorem to rough paths*, 2007, pp. 835–847. MR2348055

[Lyo98] T. J. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 215–310. MR1654527

[Mél96] S. Méléard, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), 1996, pp. 42–95. MR1431299

[Mao08] X. Mao, *Stochastic differential equations and applications*, Horwood, 2008.

[McK66] H. P. McKean Jr., *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Nat. Acad. Sci. U.S.A. **56** (1966), 1907–1911. MR221595

[McK67] H. P. McKean Jr., *Propagation of chaos for a class of non-linear parabolic equations*, Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967), 1967, pp. 41–57. MR0233437

[MPR17] T. Mastrolia, D. Possamaï, and A. Réveillac, *On the Malliavin differentiability of BSDEs*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 464–492. MR3606749

[MSS12] C. J. Mode, C. K. Sleeman, and W. Scientific, *Stochastic processes in genetics and evolution: Computer experiments in the quantification of mutation and selection*, World Scientific Publishing Company, 2012.

[MSS94] A. Millet and M. Sanz-Solé, *A simple proof of the support theorem for diffusion processes* **1583** (1994), 36–48. MR1329099

[MV16] Y. S. Mishura and A. Y. Veretennikov, *Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations*, arXiv preprint arXiv:1603.02212 (2016).

[Nua06] D. Nualart, *The Malliavin calculus and related topics*, Second, Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR2200233

[Pis89] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge Tracts in Mathematics, vol. 94, Cambridge University Press, Cambridge, 1989. MR1036275

[Pit77] L. D. Pitt, *A Gaussian correlation inequality for symmetric convex sets*, Ann. Probability **5** (1977), no. 3, 470–474. MR448705

[Pop01] S. Pope, *Turbulent flows*, IOP Publishing, 2001.

[PS11] G. Pagès and A. Sellami, *Convergence of multi-dimensional quantized SDE's*, Séminaire de Probabilités XLIII, 2011, pp. 269–307. MR2790377

[Rie17] S. Riedel, *Transportation-cost inequalities for diffusions driven by Gaussian processes*, Electron. J. Probab. **22** (2017), Paper No. 24, 26. MR3622894

[Roy14] T. Royen, *A simple proof of the gaussian correlation conjecture extended to multivariate gamma distributions*, arXiv preprint arXiv:1408.1028 (2014).

[RS17] S. Riedel and M. Scheutzow, *Rough differential equations with unbounded drift term*, Journal of Differential Equations **262** (2017), no. 1, 283 –312.

[RW18] P. Ren and J.-L. Wu, *Least squares estimation for path-distribution dependent stochastic differential equations*, arXiv preprint arXiv:1802.00820 (2018).

[Šid68] Z. Šidák, *On multivariate normal probabilities of rectangles: Their dependence on correlations*, Ann. Math. Statist. **39** (1968), 1425–1434. MR230403

[Sal20] W. Salkeld, *Small ball probabilities, metric entropy and gaussian rough paths*, arXiv preprint arXiv:2002.06447 (2020).

[San61] I. N. Sanov, *On the probability of large deviations of random variables*, Select. Transl. Math. Statist. and Probability, Vol. 1, 1961, pp. 213–244. MR0116378

[Sch87] M. Scheutzow, *Uniqueness and non-uniqueness of solutions of vlasov-mckean equations*, Journal of the Australian Mathematical Society **43** (1987), no. 2, 246–256.

[Sto96] W. Stolz, *Some small ball probabilities for Gaussian processes under nonuniform norms*, J. Theoret. Probab. **9** (1996), no. 3, 613–630. MR1400590

[Str64] V. Strassen, *An invariance principle for the law of the iterated logarithm*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **3** (1964), 211–226 (1964). MR175194

[Sug85] H. Sugita, *On a characterization of the Sobolev spaces over an abstract Wiener space*, J. Math. Kyoto Univ. **25** (1985), no. 4, 717–725. MR810975

[SV72] D. W. Stroock and S. R. S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, 1972, pp. 333–359. MR0400425

[Szn91] A.-S. Sznitman, *Topics in propagation of chaos*, École d'Été de Probabilités de Saint-Flour XIX—1989, 1991, pp. 165–251. MR1108185

[TZ15] M. Tahmasebi and S. Zamani, *Weak differentiability of solutions to SDEs with semi-monotone drifts*, Bulletin of the Iranian Mathematical Society **44** (2015), no. 4.

[ÜZ00] A. S. Üstünel and M. Zakai, *Transformation of measure on Wiener space*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000. MR1736980

[Var66] S. R. S. Varadhan, *Asymptotic probabilities and differential equations*, Comm. Pure Appl. Math. **19** (1966), 261–286. MR203230

[Var84] S. R. S. Varadhan, *Large deviations and applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 46, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984. MR758258

[vdVvZ07] A. van der Vaart and H. van Zanten, *Bayesian inference with rescaled Gaussian process priors*, Electron. J. Stat. **1** (2007), 433–448. MR2357712

[vdVvZ11] A. van der Vaart and H. van Zanten, *Information rates of nonparametric Gaussian process methods*, J. Mach. Learn. Res. **12** (2011), 2095–2119. MR2819028

[Vil09] C. Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009. Old and new. MR2459454

[WSS01] R. C. Williamson, A. J. Smola, and B. Schölkopf, *Generalization performance of regularization networks and support vector machines via entropy numbers of compact operators*, IEEE Trans. Inform. Theory **47** (2001), no. 6, 2516–2532. MR1873936

[Zab00] J. Zabczyk, *Stochastic invariance and consistency of financial models*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **11** (2000), no. 2, 67–80. MR1797512

[Zha16] X. Zhang, *Stochastic differential equations with Sobolev diffusion and singular drift and applications*, Ann. Appl. Probab. **26** (2016), no. 5, 2697–2732. MR3563191